


1948

# Polygenic inheritance of fruit size in red pepper (*Capsicum frutescens* L)

Ian Khambanonda  
*Iowa State College*

Follow this and additional works at: <https://lib.dr.iastate.edu/rtd>

 Part of the [Agricultural Science Commons](#), [Agriculture Commons](#), [Agronomy and Crop Sciences Commons](#), [Genetics Commons](#), and the [Plant Pathology Commons](#)

---

## Recommended Citation

Khambanonda, Ian, "Polygenic inheritance of fruit size in red pepper (*Capsicum frutescens* L) " (1948). *Retrospective Theses and Dissertations*. 13612.  
<https://lib.dr.iastate.edu/rtd/13612>

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact [digirep@iastate.edu](mailto:digirep@iastate.edu).

# NOTE TO USERS

This reproduction is the best copy available.

**UMI<sup>®</sup>**



POLYGENIC INHERITANCE OF FRUIT SIZE IN RED PEPPER  
(CAPSICUM FRUTESCENS L.)

by

Ian Khambanonda

A Thesis Submitted to the Graduate Faculty  
for the Degree of  
DOCTOR OF PHILOSOPHY

Major Subject: Genetics

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

Head of Major Department

Signature was redacted for privacy.

Dean of Graduate College

Iowa State College  
1948

UMI Number: DP12801

## INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

**UMI<sup>®</sup>**

---

UMI Microform DP12801

Copyright 2005 by ProQuest Information and Learning Company.

All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

ProQuest Information and Learning Company  
300 North Zeeb Road  
P.O. Box 1346  
Ann Arbor, MI 48106-1346

QH423  
K527p

TABLE OF CONTENTS

|                                 | Page |
|---------------------------------|------|
| I. INTRODUCTION.....            | 1    |
| II. MATERIALS AND METHODS.....  | 4    |
| III. GENETIC THEOREMS.....      | 8    |
| A. The Means.....               | 8    |
| B. The Variances.....           | 13   |
| C. Number of Genes.....         | 22   |
| IV. RESULTS AND DISCUSSION..... | 27   |
| A. Fruit Length and Width.....  | 27   |
| B. Fruit Shape.....             | 33   |
| C. Green Weight of Fruit.....   | 43   |
| D. Dry Weight of Fruit.....     | 53   |
| V. CONCLUSIONS.....             | 64   |
| VI. SUMMARY.....                | 68   |
| VII. ACKNOWLEDGEMENT.....       | 71   |
| VIII. LITERATURE CITED.....     | 72   |

T8680

## I: INTRODUCTION

Polygenic or quantitative characters are defined as those measurable characters which are controlled by a large number of genes usually with small individual effects. Among them are expressions such as yield and maturity, with which plant and animal breeders are chiefly concerned.) Inasmuch as the success of breeding depends largely on the inheritance of such characters, an understanding of polygenic actions would be invaluable guidance to breeders in choosing appropriate experimental techniques.

Since individual effects of polygenes are small, they are obscured by environmental variation. In a few simple cases in plants, in which they are recognizable, the genes have been identified and their nature ascertained in studies of their linkage relations with qualitative genes (Sax, 1924; Lindstrom, 1927a), of pleiotropy which may be conceived as complete linkage (De Haan, 1931; Yeager, 1937), and of direct classification of phenotypes (Powers, 1934). Although the pattern of inheritance has been shown therein to be similar to that of qualitative characters, it is still an a priori assumption that it is generally so, particularly that of a more complex character.)

Previous studies of the red pepper (Webber, 1912; Halsted, 1915; Dale, 1929; Deshpande, 1933; Kaiser, 1935; Khambanonda,

1941) have shown that measurements of fruit length, width, shape, and weight may be considered typical polygenic characters of varying degrees of complexity. A thorough genetic analysis of them will serve as an illustration of methods by which actions of the polygenes can be determined.

As the number of genes affecting a polygenic character is large, it is impossible to single them out individually. A suggestion for simplification would be to study them in small groups if grouping is feasible on biological bases. Powers (1941) follows this method of approach in his studies of inheritance of maturity in the tomato by using three developmental stages and investigating them separately. MacArthur and Butler (1938) propose that fruit-size genes be divided at least into two groups: those governing rate of cell division or duration of active mitosis, and the others governing cell expansion.

The usual genetic analysis of complex polygenic characters utilizes statistical methods by which genes and their combined effects are investigated as a whole. It involves fitting data to one model of gene action and disproving other alternatives.) The disproof is cumbersome, is frequently disregarded, and so deductions are inconclusive.

In most instances in the past, materials were limited to parental,  $F_1$ ,  $F_2$ , backcross generations, and selected advanced progenies. Often, results were inadequate for



precise conclusion. For the present study, it was thought that unselected  $F_3$  and  $F_4$  generations might supply sufficient information, and the experiment was planned accordingly.

The objects of this experiment were to determine the pattern of inheritance of fruit-size characters, length, width, shape, and weight, by successive selfings of a hybrid between two red pepper varieties which showed large differences.

(1) To find the genetic relationship of size components, whether inherited separately or in conjunction with each other. In this respect, it would be desirable to ascertain if shape genes are transmitted as such or if they are composites of length and width factors.

(2) To observe the trend of means brought about by self-fertilization and to make deductions on actions of genes therefrom.

(3) To estimate environmental, genotypic, and genetic variances in segregating generations, and to discover how these estimates conform with the expected values set forth by genetic theorems, so that inference may be made on properties and behavior of genes.

(4) To fit results to formulae for the calculation of minimum number of genic differences in materials of the cross.

## II. MATERIALS AND METHODS

Two commercial varieties of red pepper (Capsicum frutescens L.) were obtained from W. Atlee Burpee Company of Philadelphia; Red Chili ( $P_1$ ) which has small elongated fruit and Sunnybrook ( $P_2$ ) which has large oblate fruit. They were naturally self-fertilized and were uniform in fruit and other morphological characteristics.

Plants were selfed once in 1940. Crosses were then made between the two varieties, using Red Chili as the female parent.  $F_1$  plants were selfed for  $F_2$ .

Seeds from sixty unselected  $F_2$  plants were separately collected for sixty  $F_3$  progenies. Four plants of each  $F_3$  progeny were picked at random at the seedling stage and seeds were harvested and bulked by progeny for the  $F_4$  field test. The production of seed was done in the greenhouse.

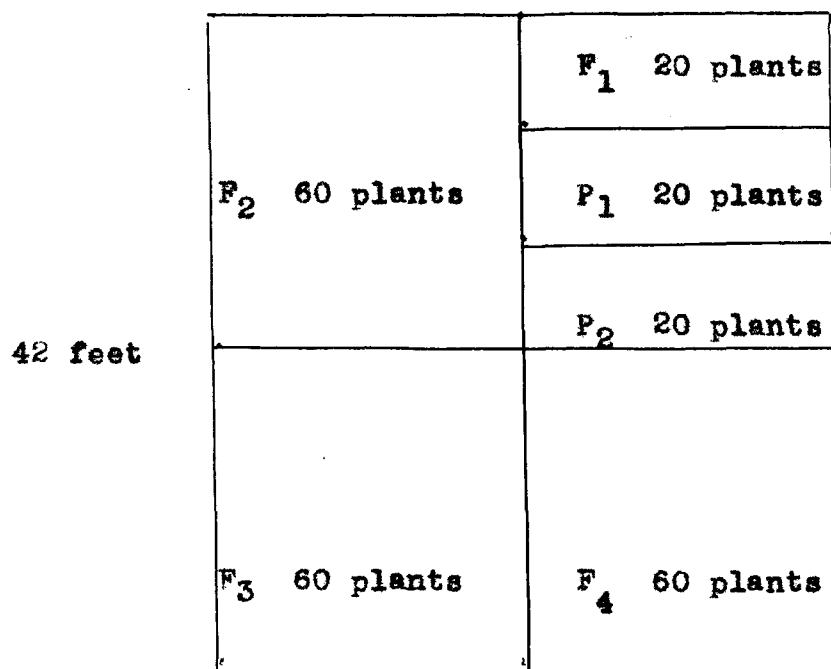
In the 1946 test, plants of all generations,  $P_1$ ,  $P_2$ ,  $F_1$ ,  $F_2$ ,  $F_3$ , and  $F_4$  were started in the greenhouse and transplanted to the field. No selection was practised. The few missing hills, about eight percent of the total of 3600 plants, were filled by replantings.

The design of experiment was a randomised complete block. Four plots were randomised within a block: one plot with sixty  $F_2$  plants; another plot with sixty  $F_3$  plants,

one plant per progeny; the third with sixty  $F_4$  plants, one plant per progeny; and the fourth with a subplot of twenty  $P_1$  plants, another subplot with twenty  $P_2$ , and the third subplot with twenty  $F_1$ . Plants were randomised within plots and subplots; and subplots were randomised within a plot.  $F_3$  and  $F_4$  plants were labeled according to the progeny to which they belonged.

The arrangement in a plot was ten plants per row spaced  $1\frac{3}{4}$  feet within the row and  $3\frac{1}{2}$  feet between rows, six rows per plot or two rows per subplot:

35 feet



BLOCK I

Fifteen replications gave a total of 3,600 hills, of which seven plants were missing. Border rows were used.

With sixty plants to a plot, one-sixtieth or 1.67 percent of the genetic variance of  $F_2$ ,  $F_3$ , and  $F_4$  was confounded with block differences and might be considered negligible (cf. 100-plant plot by Powers, 1942). Additional plants would have increased environmental variations or error within plot but would have decreased the proportion of genetic variance confounded with block. Smaller plot size would have had reverse effects.

Sixty progenies of  $F_3$  and  $F_4$  were considered as fair samples of  $F_3$  and  $F_4$  populations respectively. However, one plant per progeny per generation of  $F_3$  and  $F_4$  in a plot, completely confounded variance within progenies with block variations. More plants of fewer progenies, for example, two of each of thirty progenies, would still constitute a small sample, and because the two plants were samples of a segregating population, the effect of confounding would still be large.

Since  $P_1$ ,  $P_2$ , and  $F_1$  were non-segregating, twenty plants of each in a subplot were considered sufficient. The subplots were so arranged that comparisons of  $P_1$ ,  $P_2$ , and  $F_1$  were precise among themselves.

Five fruits from each plant were harvested when ripe or turning red. Length was measured from proximal to distal

end of individual fruits with peduncle removed, and width at middle portion, to the nearest millimeter. The five measurements were averaged. The shape index was expressed as a ratio of length to width.

Dry and green weights of five fruits per plant were recorded to the nearest tenth of a gram and to the nearest gram respectively. The total weight of five fruits per plant is presented in the results, and not the average fruit weight.

### III. GENETIC THEOREMS

#### A. The Means

Ordinarily, the method of distinguishing multiplicative action from dominance of small value of a gene is based on means.

Let  $\bar{P}_1$ ,  $\bar{P}_2$ , and  $\bar{F}_1$  be the means of one homozygous parent, of the other homozygous parent, and of first generation hybrids respectively.

$$\text{Arithmetic dominance} = \bar{F}_1 - \frac{\bar{P}_1 + \bar{P}_2}{2}$$

$$\text{Geometric dominance} = \log \bar{F}_1 - \frac{\log \bar{P}_1 + \log \bar{P}_2}{2}$$

$$F_2 \text{ arithmetic mean} = \frac{\bar{P}_1 + \bar{P}_2 + 2\bar{F}_1}{4} \text{ (Wright, 1922), which}$$

may be extended to:

$$\log F_2 \text{ geometric mean} = \frac{\log \bar{P}_1 + \log \bar{P}_2 + 2 \log \bar{F}_1}{4}$$

The calculations are dependent on values of the homozygotes ( $P_1$ ,  $P_2$ ), the heterozygote ( $F_1$ ) and the segregating population ( $F_2$ ). It may be argued that genetic backgrounds or substrates of  $P_1$ ,  $P_2$ ,  $F_1$ , and  $F_2$  are all different and uncomparable with each other. These substrate factors may modify expressivity of genes. If they do so, conclusions derived from  $P_1$ ,  $P_2$ ,  $F_1$ , and  $F_2$  data may be misleading, whereas  $F_2$ ,  $F_3$ , and  $F_4$  results should be preferred since individuals of the latter generations have comparatively similar samples of the genetic substrates.

Two criteria for distinguishing effects of genes may be presented:

(1) Variations within  $P_1$ ,  $P_2$ , and  $F_1$  indicate environmental influence on the character. If the variations in all three populations are approximately equal, the effects of environment are additive. If they become equal only by transforming data to a logarithmic scale, then the effects of environment, and probably but not necessarily those of genes, are multiplicative. As effects of genes and environment are often unparallel, the test is not conclusive.

The skewness of distributions within  $P_1$ ,  $P_2$ , and  $F_1$  indicates whether environmental variations are additive or multiplicative. However, the third-moment statistic, especially of distributions of segregating generations, may be exaggerated by infrequent inclusion of extreme individuals so that it does not always measure skewness as such.

(2) A more effective method of differentiating genic manifestations is the examination of the trend of means of successively selfed generations after hybridization.

Ignoring various genetic substrates and assuming no epistasy, suppose a character is affected by two pairs of genes:

$$\begin{array}{lll} AA = 2x_1 & Aa = x_1 / d_1 & aa = 0 \\ BB = 2x_2 & Bb = x_2 / d_2 & bb = 0 \end{array}$$

Symbols  $x_1$ ,  $x_2$ ,  $d_1$ , and  $d_2$  may represent any real values. If  $d_1$  and  $d_2$  are positive, dominance for the large value is indicated; if  $d_1$  and  $d_2$  are negative, dominance for the small value is indicated. If  $d_1$  is positive and  $d_2$  negative, both dominances are present; if  $d_1$  and  $d_2$  are zero, dominance is absent.

$$\text{Let } P_1 = AAbb \text{ and } P_2 = aaBB$$

$$\text{Mean } P_1 = 2x_1 \text{ and mean } P_2 = 2x_2$$

$$\frac{1}{2}(\text{mean } P_1 + \text{mean } P_2) = x_1 + x_2$$

$$\text{Mean } F_1 = x_1 + x_2 + d_1 + d_2$$

$$\text{Mean } F_2 = x_1 + x_2 + \frac{1}{2}(d_1 + d_2)$$

$$\text{Mean } F_3 = x_1 + x_2 + \frac{1}{4}(d_1 + d_2)$$

$$\text{Mean } F_4 = x_1 + x_2 + \frac{1}{8}(d_1 + d_2)$$

$$\text{Mean } F_{\infty} = x_1 + x_2$$

The formulae may be extended to  $r$  factors and  $n$  generations of self-fertilization:

$$\text{Mean } F_n = \sum_{i=1}^r x_i + \frac{1}{2^n - 1} \left( \sum_{i=1}^r d_i \right)$$

It may be proved that when there is no selection, linkage does not change the means of generations. Although it alters genotypic ratios and variances, the means remain the same as if linkage is absent.

By continued selfing without selection, the mean of any one generation approaches the mean of parents by one-half



of the difference between the mean of the preceding generation and the mean of parents:

$$\bar{F}_{n-1} - \bar{F}_n = \frac{1}{2}(\bar{F}_{n-1} - \frac{\bar{P}_1 + \bar{P}_2}{2}), \text{ where } n \geq 2, \text{ generations of selfing.}$$

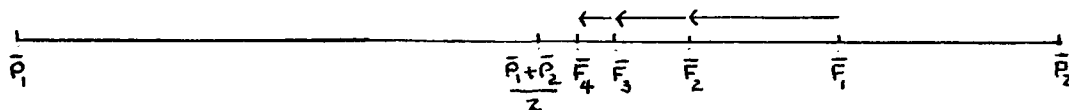
$$\text{Example: } \bar{F}_2 - \bar{F}_3 = \frac{1}{2}(\bar{F}_2 - \frac{\bar{P}_1 + \bar{P}_2}{2})$$

$$\frac{\bar{F}_{n-1} - \bar{F}_n}{\bar{F}_{n-2} - \bar{F}_{n-1}} = \frac{1}{2}, \text{ where } n \geq 3, \text{ generations of selfing,}$$

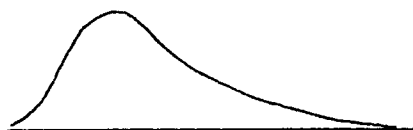
and limit  $\bar{F}_n = \frac{1}{2}(\bar{P}_1 + \bar{P}_2)$   
 $n \rightarrow \infty$

$$\text{Example: } \bar{F}_3 - \bar{F}_4 = \frac{1}{2}(\bar{F}_2 - \bar{F}_3)$$

If  $\sum d_i$  is positive (excess of dominance of large), even if some  $d_i$  are negative, the trend of generation means is decreasing and may be represented as follows:



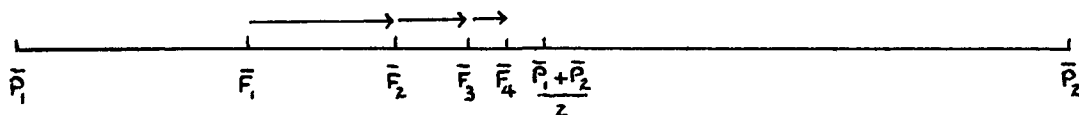
A common problem in analysis is the interpretation of positive skewness of  $F_2$  distributions, which may be explained



$F_2$  distribution

by either multiplicative effects of genes or dominance of small values.

If dominance is predominantly for small ( $\sum d_i$  negative), the trend of means of successively selfed generations is invariably increasing:



If it is not increasing, dominance of small cannot be the cause of skewness. An alternative explanation, the multiplicative effects of genes can be considered; and the original data should be transformed to logarithms for analysis.

The above formulae hold true for all degrees of dominance, partial, complete, or overdominance, positive or negative direction, or combinations of them. The sum of dominant deviations,  $\sum d_i$ , may be estimated from means of all generations.

As a corollary, if  $\sum d_i$  is zero, in which dominance of small and large is balanced, or dominance is absent, the average of parental means equals the mean of  $F_1, F_2, F_3, \dots, F_n$ . Selfing without selection does not increase or decrease the mean of any generation.

$$\frac{\bar{P}_1 + \bar{P}_2}{2} = \bar{F}_1 = \bar{F}_2 = \bar{F}_3 = \dots = \bar{F}_n$$

Epistasy or interaction of non-allelic genes would shift the means in one direction or the other essentially in the same manner as dominance, which is interaction of allelic genes. In inhibiting interaction, which is comparable to dominance of small, the means of successively selfed generations would tend to increase; whereas in complementary action, which is comparable to dominance of large, they would decrease. The rate of shift depends on the kind and magnitude of epistasy in question; and no generalisation can be formulated.

### B. The Variances

Variances of  $P_1$ ,  $P_2$ , and  $F_1$ , which are entirely environmental, may be equal. Powers (1942) finds that, in the tomato, they are proportional to generation means. Charles and Smith (1939) believe that standard deviations are correlated with means. Gustafsson (1946) finds  $F_1$  more variable than  $P_1$  and  $P_2$ , whereas it is a common experience among maize breeders that in some characters as date of tasseling, inbreds are more affected by climatic changes than  $F_1$ . In cotton, Panse (1940) does not find any relationship between variances and means of staple length.

If the environmental variance of segregating generations can be estimated from  $P_1$ ,  $P_2$ , and  $F_1$ , then genotypic variance ( $H$ ) of  $F_2$ ,  $F_3$ , and  $F_4$  is the difference between total variance ( $T$ ) and environmental variance ( $E$ ). Genotypic variance ( $H$ ) may be broken down into additively genetic variance ( $G$ ) plus variances ascribed to dominance ( $D$ ), epistasy ( $I$ ), and linkage ( $L$ ).

$$T = H + E = G + D + I + L + E$$

The estimate of genetic variance ( $G$ ) is the regression of  $F_3$  progenies on  $F_2$  parents (Panse, 1940), which measures heritability of  $F_2$ . But the estimation of genetic variance of  $F_3$  from genetic variance of  $F_2$  grown in the previous year

is not entirely correct unless genotypic manifestations of  $F_2$  and  $F_3$  are similarly affected by environment from year to year. In a majority of agronomic characters, it has been found that interaction between genotypes and year, or covariance (HE), is often significantly large. The interaction may be eliminated by use of  $F_3$  and  $F_4$  planted at the same time. Heritability of  $F_3$  is measured by regression of  $F_4$  on  $F_3$  progenies.

When homozygous parents are used, the genotypic variance of non-segregating generations is zero:

$$H(P_1) = H(P_2) = H(F_1) = 0$$

With selfing, the genotypic variances of segregating generations may be formulated as follows:

If dominant homozygote AA is  $2x$ , heterozygote Aa is  $x + d$ , and recessive homozygote aa is 0,

$$\text{Variance of } F_2 \text{ is } \frac{x^2}{2} + \frac{d^2}{4}$$

$$\text{Variance of } F_3 \text{ is } \frac{3x^2}{4} + \frac{3d^2}{16}$$

$$\text{Variance of } F_4 \text{ is } \frac{7x^2}{8} + \frac{7d^2}{64}$$

$$\text{Variance of } F_5 \text{ is } \frac{15x^2}{16} + \frac{15d^2}{256}$$

$$\text{Variance of } F_n \text{ is } [2^{n-1} - 1] \left[ \frac{x^2}{2^{n-1}} + \frac{d^2}{4^{n-1}} \right], \text{ where } n \geq 2, \text{ generations of selfing.}$$

The formulae may be extended to  $r$  independent genes with no epistatic effects:

$$H(F_n) = [2^{n-1} - 1] \sum_{i=1}^r \left[ \frac{x_i^2}{2^{n-1}} + \frac{d_i^2}{4^{n-1}} \right]$$

(1) With no epistasy or linkage,

(a) If dominance is absent,  $d = 0$ , the heterozygote being mid-way between parental homozygotes,

$$\frac{G(F_n) - G(F_{n-1})}{G(F_{n-1}) - G(F_{n-2})} = \frac{1}{2}; \text{ where } n \geq 3, \text{ generations of selfing}$$

$$G(F_3) = 1.50 G(F_2)$$

$$G(F_4) = 1.75 G(F_2)$$

$$G(F_5) = 1.875 G(F_2)$$

(b). If all factors are completely dominant, for large ( $d = x$ ) or for small ( $d = -x$ ), or both for large and small,

$$\frac{H(F_n) - H(F_{n-1})}{H(F_{n-1}) - H(F_{n-2})} = \frac{1}{4}$$

$$H(F_3) = 1.25 H(F_2)$$

$$H(F_4) = 1.3125 H(F_2)$$

$$H(F_5) = 1.328125 H(F_2)$$

(c) If some factors are completely dominant while others are not, or if factors are partially dominant ( $0 < d < |x|$ ), or combinations thereof,

$$\frac{1}{4} < \frac{H(F_n) - H(F_{n-1})}{H(F_{n-1}) - H(F_{n-2})} < \frac{1}{2}$$

(d) As selfing continues, overdominance ( $d > |x|$ ) of some size, such as  $d = |2x|$ , decreases genotypic variance,

$$H(F_n) < H(F_{n-1}).$$

The decrease may be roughly visualised by the fact that selfing lowers the proportion of heterozygotes, which, being overdominant, have extreme values away from the generation mean.

(2) Mean genotypic variance within  $F_3$  progenies equals one-half of  $F_2$  genotypic variance (Pense, 1940). It may be proved further that mean genotypic variance within  $F_4$  progenies is one-half of that within  $F_3$  progenies if there is no epistasy or linkage.

Suppose two factors are segregating independently. Heterozygote Aa contributes variance w; Bb, variance z; and AaBb, variance  $w/z$ .

Variance of  $F_2$  which is derived from  $F_1$ , AaBb, is  $w/z$ .

| <u><math>F_3</math> progenies derived from <math>F_2</math><br/>of following genotypes</u> | <u>Frequency</u> | <u>Variance within <math>F_3</math><br/>progenies</u> |
|--|------------------|---|
| AaBb   | $\frac{1}{4}$    | $w/z$   |
| AaBB, Aabb   | $\frac{1}{4}$    | w   |
| AABb, aaBb   | $\frac{1}{4}$    | z   |
| AABB, AAbb, aaBB, aabb   | $\frac{1}{4}$    | 0   |

Mean variance within  $F_3$  progenies =  $\frac{1}{2}(w/z)$

| <u><math>F_4</math> progenies derived from <math>F_3</math><br/>of following genotypes</u> | <u>Frequency</u> | <u>Variance within <math>F_4</math><br/>progenies</u> |
|--|------------------|---|
| AaBb   | 1/16             | $w/z$   |
| AaBB, Aabb   | 3/16             | w   |
| AABb, aaBb   | 3/16             | z   |
| AABB, AAbb, aaBB, aabb   | 9/16             | 0   |

Mean variance within  $F_4$  progenies =  $\frac{1}{4}(w/z)$

The relationship is true for any number of factor pairs with no epistasy or linkage, and it may be generalised that mean genotypic or genetic variance within progenies of any one generation beyond  $F_3$  is one-half of that within progenies of the preceding generation.

(3) From Fisher, Immer, and Tedin (1932), for a single factor difference, if recessive homozygote is 0, dominant homozygote  $2x$ , and heterozygote  $x/d$ , the formulae may be extended to  $F_4$ :

|                                      |                                      |
|--------------------------------------|--------------------------------------|
| Variance of $F_2$                    | $\frac{1}{4}(2x^2 + d^2)$            |
| Mean variance within $F_3$ progenies | $1/8(2x^2 + d^2)$                    |
| Mean variance within $F_4$ progenies | $1/16(2x^2 + d^2)$                   |
| Variance of means of $F_3$ progenies | $\frac{1}{4}(2x^2 + \frac{1}{4}d^2)$ |
| Variance of means of $F_4$ progenies | $\frac{1}{4}(2x^2 + 1/16 d^2)$       |

Variance due to dominance (D) may be estimated as:

$$\text{Variance of } F_2 - \text{Variance of means of } F_3 \text{ progenies} = \frac{3}{16} d^2$$

$$\text{Variance of means of } F_3 \text{ progenies} - \text{Variance of means of } F_4 \text{ progenies} = \frac{3}{64} d^2$$

If  $r$  pairs of genes are unlinked with no epistatic deviations, the above two equations become  $3/16 \sum_{i=1}^r d_i^2$  and  $3/64 \sum_{i=1}^r d_i^2$  respectively. And if dominance of each gene is equal in magnitude,  $d$ , the values finally become  $3/16 r d^2$  and  $3/64 r d^2$ .

If dominance is absent, the genetic variance of  $F_2$ ,  $G(F_2)$ , is equal to genetic variance of means of  $F_3$ , or  $F_4$ .

progenies.

(4) Qualitative factors have been demonstrated to exhibit several kinds of epistasy. Unless epistatic manifestations are outstanding, a gross study like the present one cannot identify them. Dominance and epistasy give similar statistical results, and no one has been able to separate their effects.

In polygenic characters, an effect ascribed to epistasy such as heterosis may well be explained by dominance, and vice versa. In crop improvement, both would be equally important expressions; it probably does not matter whether one or the other is the cause.

(5) Depending on phase, linkage has a role in variance of segregating populations. In coupling phase, it increases the variance; in repulsion phase, the reverse is true. The magnitude depends on crossover values and generations of inbreeding.

With continued selfings, the change in variance attributed to linkage is comparatively large in first few generations, especially when crossover values are high. After three or four self-fertilizations, which have rapidly brought homozygosis, the variance becomes constant as equilibrium is being reached when the proportion of gametes in coupling and repulsion phases is constant. To confirm the statement, formulae by Lindstrom (1948) may be cited:



$$r_n = 2(r_0 / s_0)r_{n-1} - (r_0 - s_0)^{n-1}$$

$$s_n = 2(r_0 / s_0)s_{n-1} / (r_0 - s_0)^{n-1}$$

where  $r:s$  is gametic ratio,  $s = 1$ , and  $n =$  generation of self-fertilization.

$$r_\infty = \frac{r_0 / 1}{r_0 / 3} \text{ at equilibrium.}$$

The equations are solved and values plotted in Figure 1. Thus, in advanced generations, in which homozygotes are prevalent, the effect of linkage on shift in variance is negligible.

With linkage, the rate of homozygosis of gene combinations may be formulated. If  $p$  is a crossover value of two linked genes and  $q = 1-p$ ,

$$\text{homozygosis of gene combinations in } F_2 = \frac{1}{2}(p^2 / q^2)$$

$$\text{homozygosis of gene combinations in } F_3 = \frac{1}{4}(p^2 / q^2)^2 / \frac{1}{2}$$

$$\text{homozygosis of gene combinations in } F_4 = 1/8(p^2/q^2)^3 / 3/4$$

$$\text{homozygosis of gene combinations in } F_n =$$

$$1 - \frac{1}{2^{n-2}} / \left[ \frac{p^2 / q^2}{2} \right]^{n-1}$$

The values are plotted in Figure 2, which shows that self-fertilization with linkage brings more rapid homozygosis of di-hybrid gene combinations than with independent inheritance.

If there is linkage, the mean variance within  $F_3$  progenies is not expected to be one-half of  $F_2$  variance.

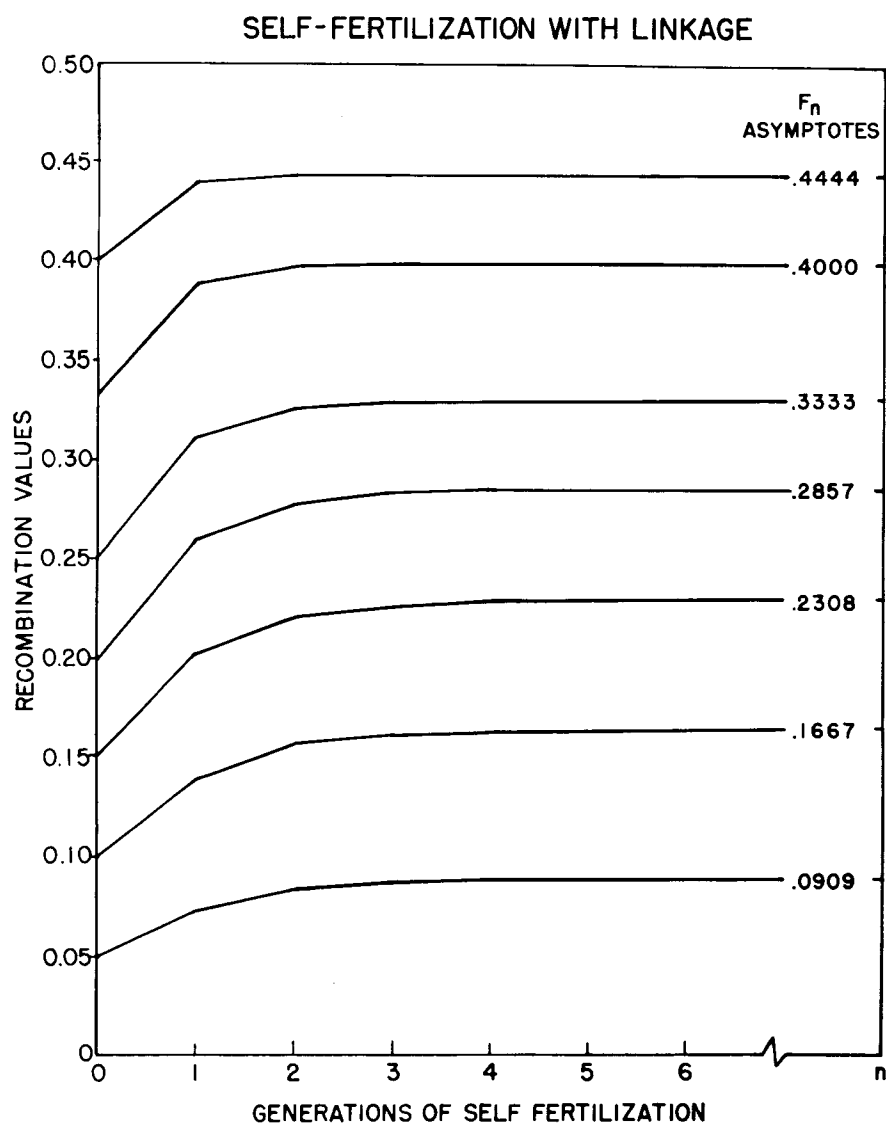


Fig. 1. Gene recombinations by self-fertilization with linkage.

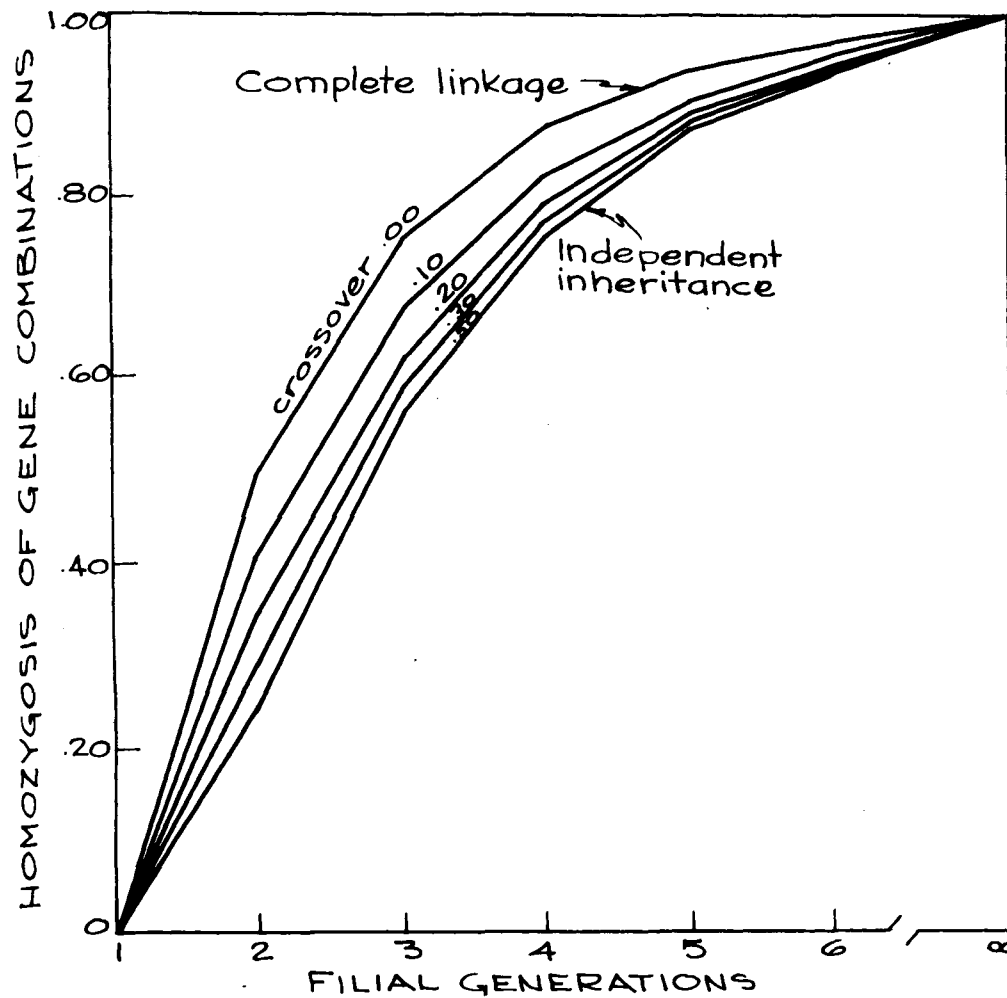


FIG.2 HOMOZYGOSIS OF GENE COMBINATIONS BY SELF-FERTILIZATION WITH LINKAGE OF A DIHYBRID.

Assuming Aa with variance w, Bb with variance z, and the two genes linked with no epistatic interaction, the variance contributed by a double heterozygote AaBb is not  $(w + z)$  but  $(w + z + \text{covariance } wz)$ , (cf. Pense, 1940). The covariance component is ascribed to linkage; it is positive in case of coupling and negative in case of repulsion. Hence, the relation, that the mean variance within  $F_3$  progenies is one-half of  $F_2$  variance, or the mean variance within  $F_4$  progenies is one-half of that within  $F_3$  progenies, is true only when linkage or epistasy is absent. The variance of  $F_3$  or  $F_4$  variances is also dependent upon intensity and phase of linkage.

### C. Number of Genes

The conventional estimation of number of genic differences in any cross is based on a ratio by arbitrary apportionment of  $F_2$  when separation of phenotypes is possible in a multimodal distribution. However, such simple cases are rare. Even then, the conclusion depends largely on how  $F_2$  individuals are classified into groups.

For unimodal distributions, several investigators have presented methods of analysis which are essentially based on the expansion of a binomial  $(x + y)^{2r}$ , where x and y are allelic gene values and r is number of gene pairs. The genes are

assumed to be independently assortive, having equal additive effects, with no dominance or epistasy.

Castle (1921) and Wright (1934) devised a formula for estimating gene numbers which may be extended. If  $r$  is number of pairs of unlinked genes with additive actions, no dominance or epistasy, and  $x_1$  is a single gene effect, then  $2 \sum_{i=1}^r x_i$  is the genetic range ( $\Delta$ ). The  $F_2$  genetic variance,  $G(F_2)$ , is  $\frac{1}{8} \sum_{i=1}^r x_i^2$ .

$$\frac{(\text{Genetic range})^2}{\text{Genetic variance of } F_2} = \frac{\Delta^2}{G(F_2)} = \frac{4 \left( \sum_{i=1}^r x_i \right)^2}{\frac{1}{8} \sum_{i=1}^r x_i^2}$$

If the genes have equal effects,  $x$ , in which a minimum number of genes is estimated, then

$$\frac{\Delta^2}{G(F_2)} = \frac{4r^2x^2}{\frac{1}{8}rx^2} = 8r$$

From previous derivations of relationship of variances of segregating generations, it may be shown that:

$$\frac{\Delta^2}{G(F_3)} = \frac{16r}{3}, \text{ and } \frac{\Delta^2}{G(F_4)} = \frac{32r}{7}$$

The difference between homozygous parental values has often been used as the genetic range. Shull (1921) points out that this is incorrect for the parents need not be extreme genotypes of the range, and, in polygenic characters, the probability of their being so is very small. If the number

of genes is large, samples of  $F_2$  also are unlikely to include the extremes.

With no linkage or epistasy, if dominance is complete in one direction only, and the genes are additive, the genetic range is  $2(\bar{P}_1 - \frac{\bar{P}_1 \bar{P}_2}{2})$ . If the dominant phenotype of a gene pair is  $2x_1$ , recessive 0, and number of gene pairs  $r$ , then the genetic range ( $\Delta$ ) is  $2 \sum_{i=1}^r x_1$ , and the  $F_2$  genotypic variance ( $H$ ) $F_2$  is  $3/4 \sum_{i=1}^r x_1^2$ .

$$\frac{(\text{Genetic range})^2}{\text{Genotypic variance of } F_2} = \frac{\Delta^2}{H(F_2)} = \frac{4(F_1 - \frac{P_1 P_2}{2})^2}{H(F_2)} = \frac{4(\sum_{i=1}^r x_1)^2}{\frac{3}{4} \sum_{i=1}^r x_1^2}$$

With equal gene effects,  $x$ ,

$$\frac{\Delta^2}{H(F_2)} = \frac{16r}{3} \text{ (F. S. Straus, unpublished)}$$

$$\frac{\Delta^2}{H(F_3)} = \frac{64r}{15}, \text{ and } \frac{\Delta^2}{H(F_4)} = \frac{256r}{63}$$

If dominance deviations are both positive and negative in directions, the genetic range will be larger than  $2(\bar{P}_1 - \frac{\bar{P}_1 \bar{P}_2}{2})$ . Selection for extreme values for several generations will give a close estimate of the range.

Analysing Winter's data in the selection of high and low oil in maize, 'Student' (1934) takes the difference between the highest and the lowest oil contents attained in 28 years as

the genetic range. He calculates additive genetic variance (G) by regression of progeny on selected parents. Since the variance is additively genetic, and is used in the estimation of minimum number of genes, he has to make many assumptions. Among them, he ignores the observed correlation between standard deviations and means, which may cause error in the genetic variance estimate. The correlation may be an explanation of the greater progress of selection for high oil content than for low oil. Also, because the original material is not an  $F_1$  but a random sample of a population, gene frequency on the average is unlikely to be one-half, which he assumes.

By a similar procedure, Panse (1940) estimates genetic variance of  $F_2$  by regression of  $F_3$  progenies on  $F_2$ , and calculates the "effective" number of factors as a ratio of squared genotypic variance within  $F_3$  progenies to variance of the genotypic  $F_3$  variance. As he cannot find genotypic variance, he substitutes genetic variance for genotypic. He proceeds to fit the results to various models of genes, some with dominance and multiplicative effects, which would contradict the definition of genetic variance, which he has actually calculated.

Assuming no linkage or epistasy, and all  $r$  factors segregating with equal variance  $w$  in  $F_2$ , Panse proves that

$$\text{Mean genotypic variance within } F_3 \text{ progenies} = \bar{H}(F_3) = \frac{rw}{2}$$

Variance of genotypic variance within  $F_3$  progenies =

$$V[\bar{H}(F_3)] = \frac{rw^2}{4}$$

$$\frac{[\bar{H}(F_3)]^2}{V[\bar{H}(F_3)]} = \frac{4r^2w^2}{4r w^2} = r = \text{"effective" number of genes.}$$

For  $F_4$  progenies, it may be proved that:

$$\text{Mean genotypic variance within } F_4 \text{ progenies} = \bar{H}(F_4) = \frac{rw}{4}$$

Variance of genotypic variance within  $F_4$  progenies =

$$V[\bar{H}(F_4)] = \frac{3rw^2}{16}$$

$$\frac{[\bar{H}(F_4)]^2}{V[\bar{H}(F_4)]} = \frac{r^2w^2}{3rw^2} = \frac{r}{3}$$

The formulae also apply with the use of genetic variances:

$$\frac{[\bar{G}(F_3)]^2}{V[\bar{G}(F_3)]} = r, \text{ and } \frac{[\bar{G}(F_4)]^2}{V[\bar{G}(F_4)]} = \frac{r}{3}$$

When factors have multiplicative effects, transformation of data to logarithms will be sufficient and simple with the use of punched cards. For estimating number of genes, Poole and Grimball (1945) present a method of geometric analysis, whereby data need not be transformed, but complications involved are unjustified.



#### IV. RESULTS AND DISCUSSION

##### A. Fruit Length and Width

Frequency distributions of fruit length and width in all generations are presented in Figures 3 and 4, the means and variances in Table 1, and correlation coefficients in Table 2.

For length, the  $F_1$  mean is larger than that of either parent. The  $F_2$ ,  $F_3$ , and  $F_4$  distributions are multimodal and positively skewed, and are transgressive so far beyond parental and  $F_1$  values that length alone can hardly be genetically explicable without invoking extremely exceptional cases. The extensive segregation does not fit any generally known model of qualitative factors. However, it is unlikely that genes are controlling length per se, but rather length is inherited as a component of, or in conjunction with, other size expressions, notably those of fruit shape.

Fruit width of  $F_1$  is intermediate between  $P_1$  and  $P_2$ . The  $F_2$ ,  $F_3$ , and  $F_4$  individuals distribute multimodally between parental values, but the distributions cannot be sharply divided into classes. The small parent  $P_1$  is recovered in the selfed generations, whereas the large width of  $P_2$  is not.

In Table 2, the correlation between length and width is positive in  $P_1$ ,  $P_2$ , and  $F_1$ , but is negative in  $F_2$ ,  $F_3$ , and  $F_4$ .

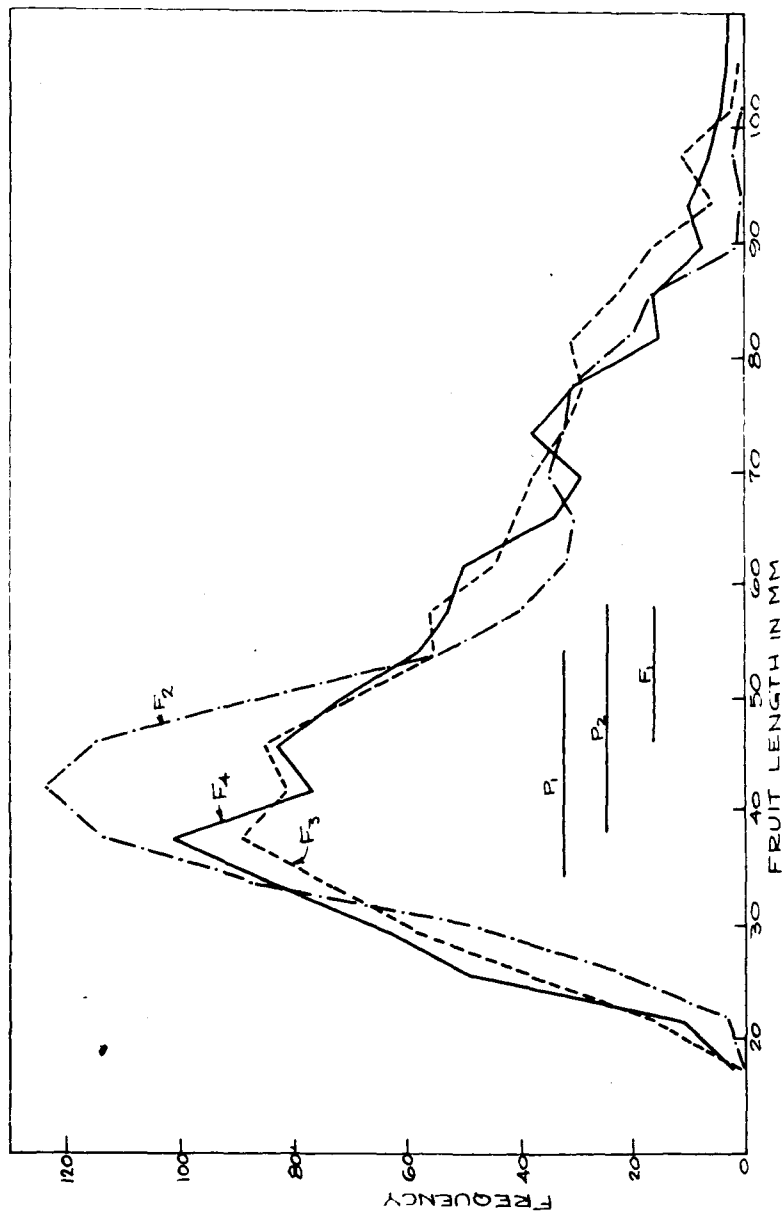


FIG. 3 FREQUENCY DISTRIBUTIONS OF FRUIT LENGTHS WITHIN GENERATIONS

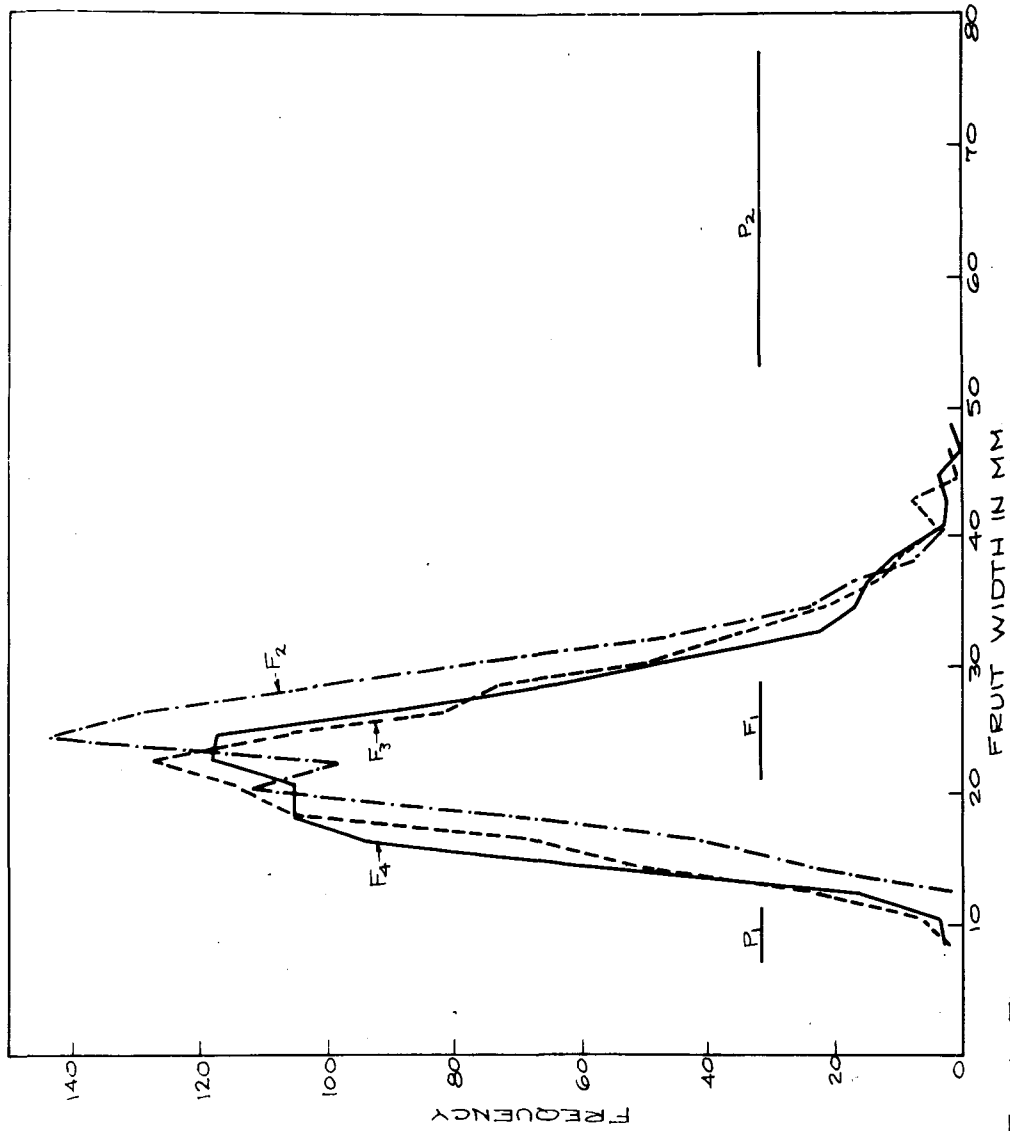


FIG. 4 FREQUENCY DISTRIBUTIONS OF FRUIT WIDTHS WITHIN GENERATIONS

Table 1. Means and variances of fruit length and width

|                             | df. | Length     |             | Width      |             |
|-----------------------------|-----|------------|-------------|------------|-------------|
|                             |     | Mean<br>mm | Mean Square | Mean<br>mm | Mean Square |
| Blocks                      | 14  |            | 1.6143      |            | 0.3214      |
| Error                       | 42  |            | 1.3481      |            | 0.1762      |
| P <sub>1</sub> (Red Chili)  | 285 | 43.5       | 0.0881      | 8.4        | 0.0056      |
| P <sub>2</sub> (Sunnybrook) | 283 | 48.2       | 0.0974      | 65.6       | 0.1078      |
| F <sub>1</sub>              | 285 | 51.3       | 0.0620      | 25.1       | 0.0246      |
| F <sub>2</sub>              | 885 | 48.9       | 2.3468      | 25.0       | 0.2717      |
| F <sub>3</sub> total        | 883 | 51.5       | 3.3746      | 23.4       | 0.4003      |
| between<br>progenies        | 59  |            | 26.3964     |            | 3.3346      |
| within<br>progenies         | 824 |            | 1.7262      |            | 0.1902      |
| F <sub>4</sub> total        | 882 | 50.1       | 3.3720      | 23.2       | 0.3855      |
| between<br>progenies        | 59  |            | 23.6297     |            | 2.8100      |
| within<br>progenies         | 823 |            | 1.9262      |            | 0.2117      |

Table 2. Correlations within generations

| *Correlation between                          | P <sub>1</sub> | P <sub>2</sub> | F <sub>1</sub> | F <sub>2</sub> | F <sub>3</sub> | F <sub>4</sub> |
|---|----------------|----------------|----------------|----------------|----------------|----------------|
| Length and width                              | .56            | .15            | .12            | -.37           | -.24           | -.26           |
| Length and width independent of green weight  | .12            | -.45           | -.11           | -.72           | -.74           | -.71           |
| Length and green weight                       | .73            | .53            | .39            | .12            | .17            | .17            |
| Length and green weight independent of width  | .57            | .64            | .39            | .68            | .73            | .69            |
| Width and green weight                        | .68            | .75            | .53            | .77            | .85            | .82            |
| Width and green weight independent of length  | .47            | .80            | .52            | .88            | .93            | .91            |
| Shape and dry weight                          | -.13           | .14            | -.28           | -.22           | -.22           | -.19           |
| Shape and dry weight (logarithmic analysis)   | -.11           | .14            | -.27           | -.20           | -.21           | -.18           |
| Shape and green weight                        | -.10           | -.03           | -.21           | -.26           | -.29           | -.26           |
| Shape and green weight (logarithmic analysis) | -.07           | .005           | -.20           | -.28           | -.31           | -.29           |

\*Significant levels for P<sub>1</sub>, P<sub>2</sub>, and F<sub>1</sub> 0.11 at 5% level,  
0.15 at 1% level.

for F<sub>2</sub>, F<sub>3</sub>, and F<sub>4</sub> 0.06 at 5% level,  
.08 at 1% level.

generations. This is one of the reasons offered by Sinnott (1935) for a theory that factors controlling shape exist. In non-segregating generations, in which variation is not genetic, any environmental deviation in one dimension, length, may be positively correlated with that in the other dimension, width. In segregating generations, if length and width genes are not linked, the genetic correlation should be zero. If they are genetically linked, the correlation, its magnitude depending on crossover values, should be higher in  $F_2$  than in  $F_3$ , and higher in  $F_3$  than in  $F_4$ . In case of complete linkage, in which length and width together express shape, a strong correlation is expected.

The length and width correlation coefficients of  $-.37$ ,  $-.24$ , and  $-.26$ , although highly significant, are of mediocre magnitude as compared with those attributed to developmental processes, such as the partial correlation between length and width independent of weight (or at constant weight), and that between length and green weight independent of width. Also, the correlation coefficients of length and green weight ( $.12$ ,  $.17$ ,  $.17$ ) in the  $F_2$ ,  $F_3$ , and  $F_4$  are small, whereas those of width and green weight ( $.77$ ,  $.85$ ,  $.82$ ) are large. Consequently, it cannot be too conclusively inferred that dimensional growths are entirely governed by shape factors. Some independent genes for length or width may exist in this cross of Red Chili and Sunnybrook.

Sharp segregations of shape indices in the  $F_2$ ,  $F_3$ , and  $F_4$  (Figures 5 and 6) give evidence that shape genes are major operative factors, which regulate the relative growths of length and width. The final fruit length and width are largely manifestations of interaction between fruit weight and shape.

Shape, in general, is negatively correlated with dry or green weights, whether analysed on an arithmetic or logarithmic scale. Although the correlations in the segregating generations are highly significant, the values (-.18 to -.31) are relatively small. If shape factors are linked with weight factors, most of them must be but loosely so.

The existence of shape genes as major determiners of relative dimensional growth rates was demonstrated in the red pepper by Kaiser (1935). He observed that the ovaries from primordium stage to anthesis were essentially similar in shape in all varieties and crosses which he investigated. After anthesis, he found two rates of relative growth in length and width, a straight line and a sickle-shaped curve. In one cross, the  $F_2$  relative-dimensional growth curves could be divided approximately into three sickle-shaped to one straight-lined, indicating a single gene pair segregating, but the mature fruit shapes did not exhibit that ratio. In the  $F_2$  of the other cross, he obtained a ratio of three straight-lined developmental curves to one sickle-shaped, dominance

being reversed, and the mature fruit shapes could be separated into two distinct groups. In other species, tomato (Yeager, 1937), squash (Sinnott and Durham, 1929), cucumber (Hutchins, 1934), and watermelon (Weetman, 1937), fruit-shape dissimilarities were shown to be fixed before anthesis, in the ovary primordia.

For the above reasons, it seems justified to analyse length and width as a ratio, or shape index, rather than to investigate them separately. Shape factors, indeed, have been identified by several investigators. Lindstrom (1927b) located a shape gene linked with "dwarf" and "peach" on first chromosome of the tomato. Nevertheless, this does not imply that factors for length or width do not exist. They may be present, but have only minor effects in the development of red pepper fruit as compared with shape factors.

#### B. Fruit Shape

The absolute values of shape indices (length divided by width) are plotted in Figure 5, and their logarithms in Figure 6. The frequencies can be combined in each generation since block variation is very small as shown in Table 3.

It is desirable first to decide whether the analysis should be made on the arithmetic or logarithmic scale. From Table 3, the trend of means of  $F_1$  to  $F_4$  is generally increasing. The  $F_1$  (2.04) is lower than the arithmetic average of  $P_1$  and



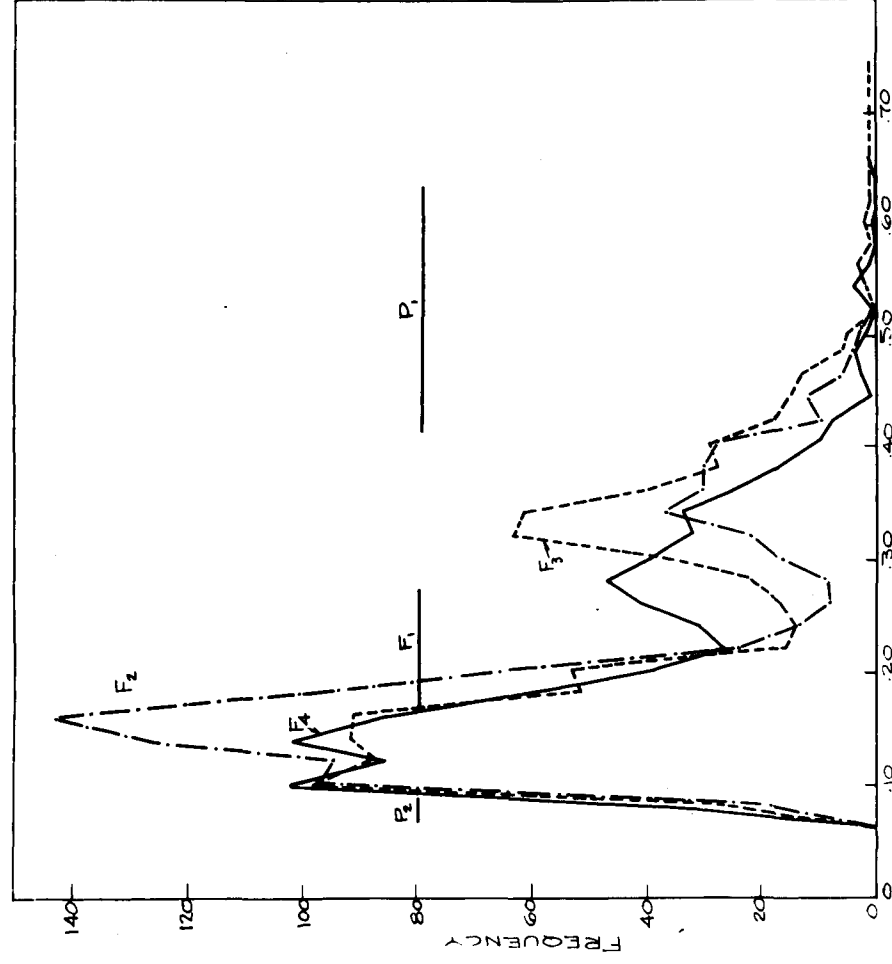


FIG. 5 FREQUENCY DISTRIBUTIONS OF FRUIT-SHAPE INDICES WITHIN GENERATIONS

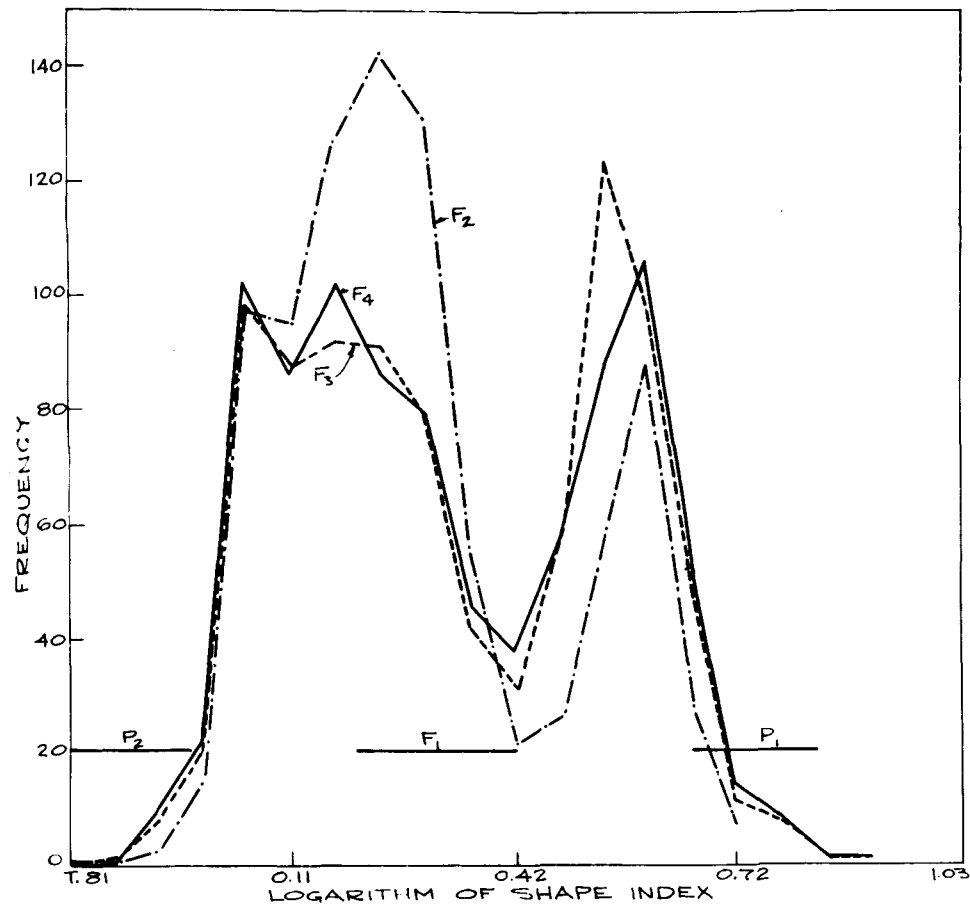


FIG. 6 FREQUENCY DISTRIBUTIONS OF LOGARITHMS OF FRUIT-SHAPE INDICES WITHIN GENERATIONS

$P_2$  (2.97). These facts, however, do not distinguish dominance of small from multiplicative actions of genes. The logarithmic means also do not serve as a criterion.

Variance of  $P_1$ , which has high shape index (elongated fruit), is the largest of all variances of non-segregating generations, that of  $F_1$ , intermediate, and that of  $P_2$  (oblate), the smallest. The ratios of variances to means follow an exponential curve. A regression line may be fitted to standard deviations and means for estimating environmental variation of  $F_2$ ,  $F_3$ , and  $F_4$ . But, by transforming data to logarithms, the variances of  $P_1$ ,  $P_2$ , and  $F_1$  are nearly equal, averaging .001111. Thus, environmental variation is multiplicative, and probably so are gene effects. Relying on this parallelism, the author decided to use logarithms; hence, multiplicative gene actions are assumed.

Logarithmic distributions of  $F_2$ ,  $F_3$ , and  $F_4$  are positively skewed, each with three modes, suggesting a major gene for fruit shape with incomplete dominance of oblate type. The homozygous dominant class cannot be isolated from the heterozygotes, but the recessives can be separated from the other two types. When the separation is made, the observed frequencies may be tested against the monohybrid ratios.

Table 3. Means and variances of fruit-shape indices

| df.                            | Arithmetic Analysis |                    |             | Logarithmic Analysis |             |                                 |
|--------------------------------|---------------------|--------------------|-------------|----------------------|-------------|---------------------------------|
|                                | Mean                | Standard Deviation | Mean Square | Mean                 | Mean Square | Genotypic <sup>2</sup> Variance |
|                                | mm.                 |                    |             | log.                 |             |                                 |
| Blocks                         | 14                  |                    | 0.7586      |                      | .022986     |                                 |
| Error                          | 42                  |                    | 0.6800      |                      | .022207     |                                 |
| P <sub>1</sub> (Red Chili)     | 285                 | 5.19               | .378        | 0.1428               | .714        | .001082                         |
| P <sub>2</sub> (Sunnybrook)    | 283                 | 0.74               | .056        | 0.0031               | -.135       | .001114                         |
| F <sub>1</sub>                 | 285                 | 2.04               | .150        | 0.0225               | .309        | .001137                         |
| F <sub>2</sub>                 | 885                 | 2.10               |             | 0.9943               | .280        | .034527                         |
| F <sub>3</sub> total           | 883                 | 2.42               |             | 1.4299               | .320        | .047091                         |
| between progenies <sup>1</sup> | 59                  |                    |             | 11.2070              |             | .377452                         |
| within progenies               | 824                 |                    |             | 0.7296               |             | .023436                         |
| F <sub>4</sub> total           | 882                 | 2.40               |             | 1.4741               | .325        | .047748                         |
| between progenies <sup>1</sup> | 59                  |                    |             | 9.0846               |             | .324955                         |
| within progenies               | 823                 |                    |             | 0.9285               |             | .027876                         |

<sup>1</sup> Variance between progenies = Variance within progenies / 15 (Genotypic variance between progeny means)

<sup>2</sup> Genotypic variance = Total variance - Environmental variance of .001111

Table 4. Estimated variances of fruit shape indices, logarithmic analysis

Average environmental variance = .001111

Regression of  $F_4$  progeny means on  $F_3$  progeny means =  
.6800

| Generations           | Genotypic Variance     | Genetic Variance       |
|-----------------------|------------------------|------------------------|
| $F_2$                 | .033416                | .017111                |
| $F_3$ total           | .045980                | 1.50(.017111)          |
| between progeny means | .023601                | .017111                |
| within progenies      | $\frac{1}{2}(.033416)$ | $\frac{1}{2}(.017111)$ |
| $F_4$ total           | .046637                | 1.75(.017111)          |
| between progeny means | .019805                | .017111                |
| within progenies      | $\frac{1}{4}(.033416)$ | $\frac{1}{4}(.017111)$ |

Genotypic variance between  $F_3$  progeny means =  $\frac{1}{15} (.377452 - .023436) = .023601$

Genotypic variance between  $F_4$  progeny means =  $\frac{1}{15} (.324955 - .027876) = .019805$

Genetic variance between  $F_3$  progeny means =  $.6800 \frac{.377452}{15} = .017111$

|                | <u>Log Class intervals</u> |                   | Expected ratio | $\chi^2$ |
|----------------|----------------------------|-------------------|----------------|----------|
|                | <u>-.22 to .40</u>         | <u>.41 to .88</u> |                |          |
| F <sub>2</sub> | 679                        | 221               | 3 : 1          | 0.10     |
| F <sub>3</sub> | 538                        | 360               | 5 : 3          | 2.51     |
| F <sub>4</sub> | 534                        | 363               | 9 : 7          | 3.81     |

The last value of chi-square (3.81) is close to 3.84, the significant level at 5% probability. This test of significance should not be too strongly emphasized for deviations from expected ratios depend much upon the partitioning of numbers in the class between the two groups. However, the segregation characteristic of a single gene pair is reasonably sharp. If modes can be used to evaluate genotypes, which is a reasonable supposition in this case, then, on a logarithmic scale, the variances may be calculated for this monogenic segregation.

|                                      |                          |
|--------------------------------------|--------------------------|
| Oblate (OO) genotype                 | = .06 (shape index 1.15) |
| Heterozygote (Oo)                    | = .18 (shape index 1.51) |
| Elongated (oo) genotype              | = .54 (shape index 3.43) |
| Genotypic variance of F <sub>2</sub> | = .0324                  |
| Genotypic variance of F <sub>3</sub> | = .0459                  |
| Genotypic variance of F <sub>4</sub> | = .0520                  |

Genotypic variances of F<sub>2</sub> and F<sub>3</sub> are slightly lower than the observed (.033416 and .045980), whereas genotypic

variance of  $F_4$  is higher than the observed (.046637), leaving no variance for any other genes. However, the dominants (00) and the recessives (oo) do not exactly correspond to the  $P_2$  mean (-.133) and the  $P_1$  mean (.714) respectively. It may be inferred that the difference between the genotype 00 (at modal value of .06) and  $P_2$  type, or that between the genotype oo (at modal value of .54) and  $P_1$  type, is due to environmental variation and to the expression of this gene in dissimilar genetic substrates.

The heterozygous value (0o) minus the average of homozygous values (00 and oo) gives the dominant deviation, d.

$$\text{Dominance of gene 0 over o} = d = .18 - \frac{.54 + .06}{2} = -.12$$

$$d^2 = .0144$$

With the observed generation means, the calculated dominant deviation is far from expected:

$$d = 4(F_2 \text{ mean} - F_3 \text{ mean}) = 4(.280 - .330) = -.200$$

$$d = 8(F_3 \text{ mean} - F_4 \text{ mean}) = 8(.330 - .325) = .040$$

By subtracting the averaged environmental variance (.001111) from the total variances of  $F_2$ ,  $F_3$ , and  $F_4$ , the genotypic variances may be computed as listed in Table 4. The ratios of genotypic variances are as follows:

$$\frac{H(F_3) - H(F_2)}{H(F_2) - H(F_1)} = \frac{.045980 - .033416}{.033416} = .38$$

$$\frac{H(F_4) - H(F_3)}{H(F_3) - H(F_2)} = \frac{.046637 - .045980}{.045980 - .033416} = .05$$

The value .38 which is between one-half and one-quarter is expected when dominance is incomplete.

$$\begin{aligned}\text{Variance of } F_2 - \text{Variance of } F_3 \text{ progeny means} &= 3/16 d^2 \\ &= .033416 - .023601 \\ d^2 &= .052347\end{aligned}$$

$$\begin{aligned}\text{Variance of } F_3 \text{ progeny means} - \text{Variance of } F_4 \text{ progeny} \\ \text{means} &= 3/64 d^2 \\ &= .023601 - .019805 \\ d^2 &= .080981\end{aligned}$$

The two values of  $d^2$  are higher than the theoretical value of .0144 estimated from genotypes. The differences may be ascribed to various causes such as the interaction of the gene with substrate factors, or the environmental variance of .001111 may be an underestimate.

From Table 4, the regression of  $F_4$  on  $F_3$  progeny means is .6800. The genetic variance between means of  $F_3$  progenies, therefore, is .6800 ( $\frac{.377452}{15}$ ) or .017111, which is also the genetic variance of  $F_2$ . Using Wright's formula, if the parental means are taken as the genetic range (.847), and if genes are unlinked with equal effects, then,

$$\begin{aligned}\text{Minimum number of genes} &= \frac{(\text{Genetic range})^2}{8(\text{Genetic variance of } F_2)} = \\ &= \frac{(.847)^2}{8(.017111)} = \text{about } 5\end{aligned}$$

Following Panse's approach, if there is no epistasy or linkage, the mean genotypic variance within  $F_3$  progenies is  $\frac{1}{2}(.033416)$  or .016708. Variance of this variance is .0003017.



The "effective" number of genes is  $\frac{.016708}{.0003017}$  or approximately 55, which is quite high. With the use of  $F_4$  values, the number is  $\frac{4(.033416)}{.0002974} \times 3$  or 84, where variance of  $F_4$  variance is .0002974.

The genotypic variance of  $F_2$  (.033416) used in the estimation is acceptable when compared with the genetic variance of  $F_2$  (.017111), which is derived from the regression of .6800. However, the variance of variance may be too low and inaccurate since, with the present experimental design, the plants within each progeny are confounded with block variation. Statistically speaking, the interaction of blocks and genotypes within progeny is significant. An evidence of this interaction is shown by the fact that the observed variance within  $F_3$  progenies is .023436 and that within  $F_4$  progenies .027876. Theoretically, the former is expected to be twice as large as the latter if there is no environmental or genetic interaction. Because of the interaction, both variances have to be estimated from  $F_2$  genotypic variance with the condition that epistasy and linkage are absent. Also, it may be added that an underestimate of environmental variance will give too high an answer for the number of genes.

### C. Green Weight of Fruit

Green weight measures the weight of five fresh ripe fruits per plant. Distributions of green weights in various generations are presented in Figures 7 and 8, and the statistics in Tables 5 and 6. Since block differences are insignificant, the frequencies can be combined in plotting the curves.

With arithmetic analysis (Figure 7 and Table 5), the distributions are unimodal, and those of  $F_2$ ,  $F_3$ , and  $F_4$  positively skewed, indicating dominance of small or multiplicative variations. The segregation, which makes the curve of a selfed generation flatter than that of the preceding one, is accompanied with increasing magnitude of kurtosis ( $g_2$ ) and variances. The small parent  $P_1$  is recovered in the segregating generations, but the large  $P_2$  is not. The  $F_1$  is closer to  $P_1$  than to  $P_2$  and it is much smaller than the average of the two parents (222.0).

Variances of  $P_1$ ,  $P_2$ , and  $F_1$  increase exponentially with means, thus the environmental variation is multiplicative. The decreasing trend of means from  $F_2$  to  $F_4$  is characteristic of gene actions with excess of dominance of large or epistasy of large. The hypothesis that dominance of small is the cause of skewness would be untenable. The alternative hypothesis, the multiplicative gene actions, may be accepted. However,

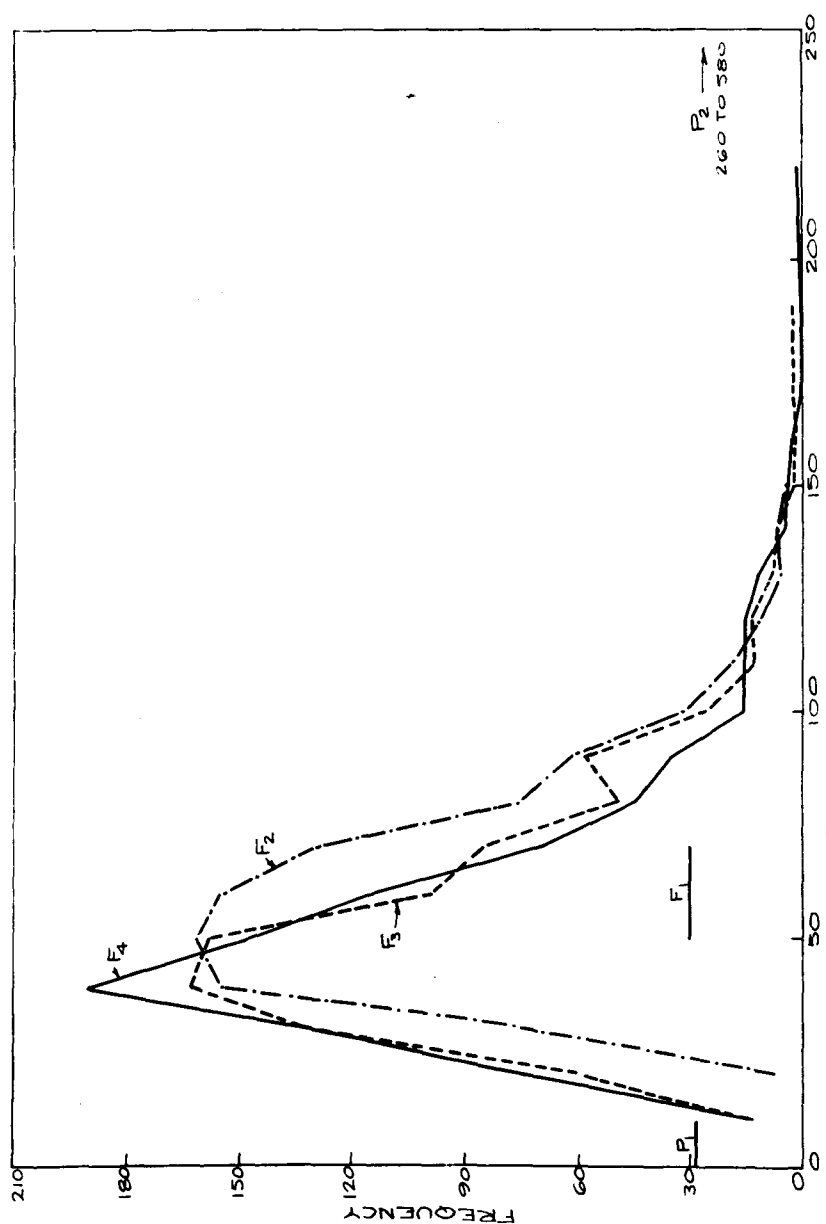


FIG. 7 FREQUENCY DISTRIBUTIONS OF GREEN FRUIT WEIGHTS WITHIN GENERATIONS

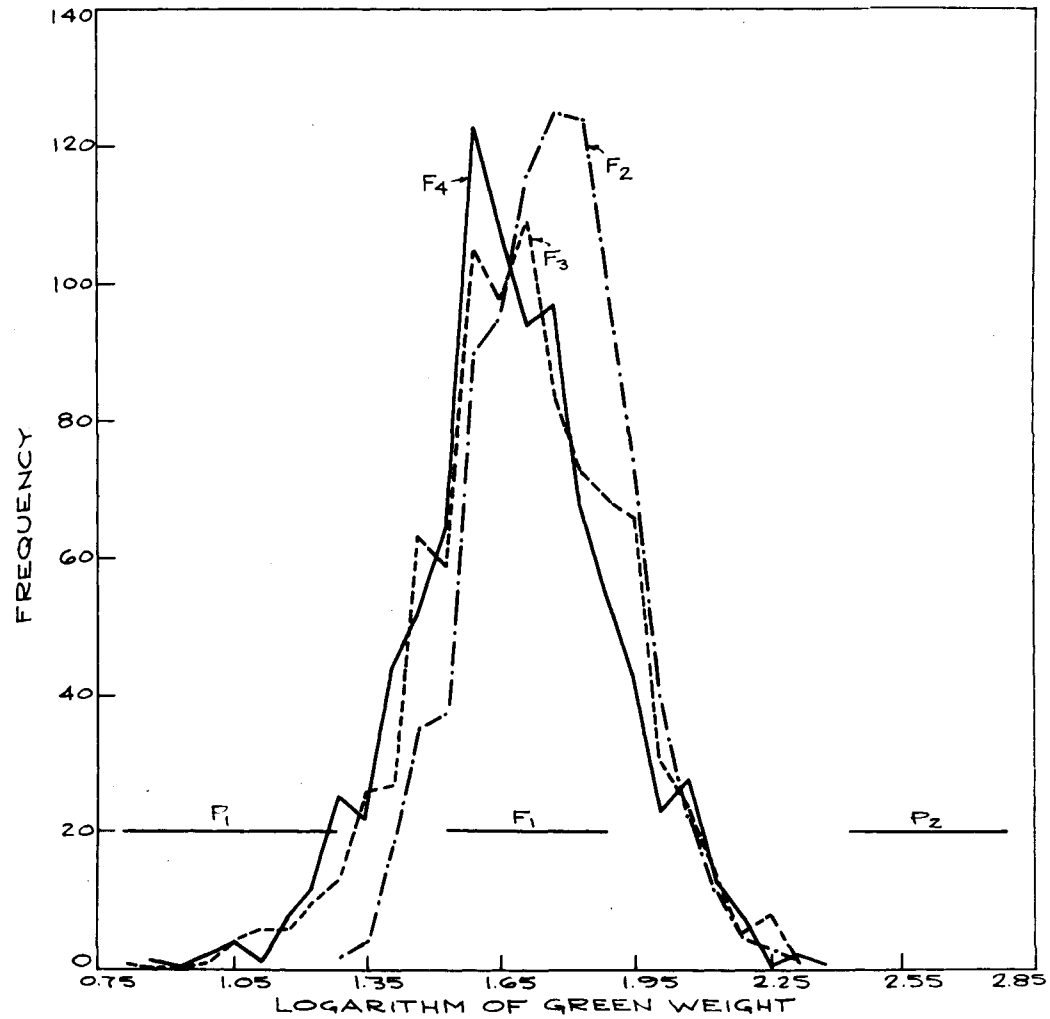


FIG. 8 FREQUENCY DISTRIBUTIONS OF LOGARITHMS OF GREEN FRUIT WEIGHTS WITHIN GENERATIONS

Table 5. Statistics of green fruit weight,  
arithmetical analysis

|                             | df. | Mean  | Skewness <sup>1</sup><br>s <sub>g1</sub> | Kurtosis <sup>2</sup><br>s <sub>g2</sub> | Mean Square |
|-----------------------------|-----|-------|--|--|-------------|
|                             |     | gm.   |  |  |             |
| Blocks                      | 14  |       |  |  | 1,520.86    |
| Error                       | 42  |       |  |  | 841.36      |
| P <sub>1</sub> (Red Chili)  | 285 | 9.0   | 0.62**                                   |  | 2.09        |
| P <sub>2</sub> (Sunnybrook) | 283 | 434.9 | -0.12                                    |  | 2,866.70    |
| F <sub>1</sub>              | 285 | 59.6  | -0.12                                    |  | 46.15       |
| F <sub>2</sub>              | 885 | 61.7  | 1.00**                                   | 1.26**                                   | 598.16      |
| F <sub>3</sub> total        | 883 | 56.3  | 1.40**                                   | 2.82**                                   | 881.34      |
| between progenies 59        |     |       |  |  | 8,004.00    |
| within progenies 824        |     |       |  |  | 371.34      |
| F <sub>4</sub> total        | 882 | 53.7  | 1.62**                                   | 3.96**                                   | 854.80      |
| between progenies 59        |     |       |  |  | 6,834.65    |
| within progenies 823        |     |       |  |  | 426.12      |

<sup>1</sup> s<sub>g1</sub> =  $\frac{1}{n} \sum (x_i - \bar{x})^3$  for P<sub>1</sub>, P<sub>2</sub>, and F<sub>1</sub>; s<sub>g1</sub> =  $\frac{1}{n} \sum (x_i - \bar{x})^3$  for F<sub>2</sub>, F<sub>3</sub> and F<sub>4</sub>

<sup>2</sup> s<sub>g2</sub> =  $\frac{1}{n} \sum (x_i - \bar{x})^4$  - .163

Table 6. Statistics of green fruit weight, logarithmic analysis

|                                | df. | Mean  | Skewness<br>$s_1$ | Kurtosis<br>$s_2$ | Mean Square | Genotypic Variance <sup>2</sup> |
|--------------------------------|-----|-------|-------------------|-------------------|-------------|---------------------------------|
| <b>log</b>                     |     |       |                   |                   |             |                                 |
| Blocks                         | 14  |       |                   |                   | .026607     |                                 |
| Error                          | 42  |       |                   |                   | .014407     |                                 |
| P <sub>1</sub> (Red Chill)     | 285 | 0.949 | -0.26             |                   | .004632     |                                 |
| P <sub>2</sub> (Sunnybrook)    | 283 | 2.634 | -0.35*            |                   | .003902     |                                 |
| F <sub>1</sub>                 | 285 | 1.771 | -0.62**           |                   | .002928     |                                 |
| F <sub>2</sub>                 | 885 | 1.758 | 0.13              | -0.19             | .026052     | .024531                         |
| F <sub>3</sub> total           | 883 | 1.694 | -0.15             | 0.24              | .050598     | .047077                         |
| between progenies <sup>1</sup> | 59  |       |                   |                   | .408239     |                                 |
| within progenies               | 824 |       |                   |                   | .024990     |                                 |
| F <sub>4</sub> total           | 882 | 1.673 | -0.07             | 0.27              | .050394     | .046773                         |
| between progenies <sup>1</sup> | 59  |       |                   |                   | .393992     |                                 |
| within progenies               | 823 |       |                   |                   | .025654     |                                 |

<sup>1</sup> Variance between progenies = Variance within progenies / 15 (Genotypic variance between progeny means)

<sup>2</sup> Genotypic variance = Total variance - Environmental variance of .003521

it is noted that the  $F_1$  mean is smaller than the  $F_2$  mean, which is contrary to the expectation when dominance is of large, but, by logarithmic transformation, the  $F_1$  becomes larger than the  $F_2$  mean as expected.

All evidences strongly point to a conclusion that the effects of genes and environment are multiplicative. It is necessary, therefore, to transform the original data into logarithms. After transformation, the frequency curves become normalized, except those of  $P_2$  and  $F_1$  which are negatively skewed. The trend of means from  $F_1$  to  $F_4$  decreases from 1.771 to 1.673; consequently, it may be inferred that there preponderantly exist genes with dominance or epistasy of large. But the shift is away from the average of parental values (1.791), instead of toward it. An explanation may be offered that the size genes are favored in the genetic substrates of parental populations. It is obvious that any conclusion drawn from  $P_1$ ,  $P_2$ , and  $F_1$  values would be a misconception. Since  $F_1$  (1.771) is smaller than  $\frac{P_1 + P_2}{2} = 1.791$ , one would assume dominance of small, which contradicts the  $F_1$ ,  $F_2$ ,  $F_3$ , and  $F_4$  results. Had not the data been transformed to logarithms, the contradiction would be much worse.

Inbreeding depression is a term often used to describe decreasing means with inbreeding. The depression may be brought about by two causes: segregation of genes with

dominance or epistasy of large directly affecting the character, and general low vigor of the individual limiting the expressivity of genes. Since the  $F_2$ ,  $F_3$ , and  $F_4$  distributions of green weights here seem to occupy the same range, the decreasing means can be attributed to the segregation of genes with dominance or epistasy of large, and not to the deterioration of the individual. The situation may be generally true in other naturally self-fertilized plants, of which the chromosomes are internally well balanced (Mather, 1943). Inbreeding does not lower the vigor of the plant; on the contrary, it is the crossing of different strains which does it since the chromosomes of hybrids of naturally self-fertilized species may be relationally unbalanced. If dominance or epistasy of large were not present, the hybrid means of the cross  $P_1 \times P_2$  would have been much smaller than the observed. The balanced condition of chromosomes would then be a definition of the genetic substrate which affects the individual as a whole. A deduction from experiments in hybridising naturally self-fertilized plants, therefore, should take substrate factors into consideration.

For purposes of analysis, it may be assumed that dominance of large causes the decreasing trend of means of generations for epistatic effects of polygenes are yet to be demonstrated. Indeed, some investigators have shown no epistatic interaction of polygenes. In Drosophila egg



production, Straus (1942) fails to find any interaction between genes on the three large chromosomes.

Ignoring epistasy, the excess of dominance of large,  $\Sigma d_1$ , in logarithms, may be calculated for predicting means of the next unselected selfed generations.

$$\Sigma d_1 = 4(F_2 \text{ mean} - F_3 \text{ mean}) = 4(1.758 - 1.694) = .256$$

$$\Sigma d_1 = 8(F_3 \text{ mean} - F_4 \text{ mean}) = 8(1.694 - 1.673) = .168$$

$$\text{Average} = \frac{.256 + .168}{2} = .212$$

Table 6 shows that  $P_1$  (.004632) is somewhat more variable than  $P_2$  (.003002) and  $F_1$  (.002928). If the average of  $P_1$ ,  $P_2$ , and  $F_1$  variances (.003521) can be used for estimating environmental variation of  $F_2$ ,  $F_3$ , and  $F_4$ , the genotypic variance then is the total variance minus the environmental variance as shown in Table 7. The ratios of genotypic variances of segregating generations are:

$$\frac{H(F_3) - H(F_2)}{H(F_2) - H(F_1)} = \frac{.047077 - .024531}{.024531} = .92$$

$$\frac{H(F_4) - H(F_3)}{H(F_3) - H(F_2)} = \frac{.046773 - .047077}{.047077 - .024531} = -.01$$

In order to explain these two ratios, certain genetic properties in addition to dominance must be invoked, such as linkage and interactions.

If linkage and epistasy are absent, and if dominant deviation is  $d_1$ ,

Table 7. Estimated variances of green weight,  
logarithmic analysis

Average environmental variance = .003521

Regression of  $F_4$  progeny means on  $F_3$  progeny means =  
.6441

| Generations           | Genotypic Variance     | Genetic variance       |
|-----------------------|------------------------|------------------------|
| $F_2$                 | .024531                | .017530                |
| $F_3$ total           | .047077                | 1.50 (.017530)         |
| between progeny means | .025550                | .017530                |
| within progenies      | $\frac{1}{2}(.024531)$ | $\frac{1}{2}(.017530)$ |
| $F_4$ total           | .046773                | 1.75(.017530)          |
| between progeny means | .024556                | .017530                |
| within progenies      | $\frac{1}{4}(.024531)$ | $\frac{1}{4}(.017530)$ |

Genotypic variance between  $F_3$  progeny means =  $\frac{1}{15} (.408239 - .024990) = .025550$

Genotypic variance between  $F_4$  progeny means =  $\frac{1}{15} (.393992 - .025654) = .024556$

Genetic variance between  $F_3$  progeny means =  $.6441 \frac{.408239}{15} = .017530$

$$\begin{aligned}\text{Variance of } F_2 - \text{Variance of } F_3 \text{ progeny means} &= 3/16 \sum d_1^2 \\ &= .024531 - .025550 \\ \sum d_1^2 &= -.005435\end{aligned}$$

$$\begin{aligned}\text{Variance of } F_3 \text{ progeny means} - \text{Variance of } F_4 \text{ progeny} \\ \text{means} &= 3/64 \sum d_1^2 \\ &= .025550 - .024556 \\ \sum d_1^2 &= .021205\end{aligned}$$

Since  $\sum d_1^2$  (-.005435) is negative, there must be some unidentified phenomena counteracting dominant deviations or decreasing the variance. But the value becomes positive (.021205), indicating that their effects become lessened in advanced generations of selfing. In the computation, the independence of genes was assumed. It is clear that linked genes in repulsion phase would give smaller variance than the same genes unlinked, and linkage effects slowly disappear as selfing continues. Thus, linkage, and also certain kinds of epistasy, may be used for explaining the observed results.

Heritability, as measured by regression of  $F_4$  on  $F_3$ , is .6441. The genetic variance between  $F_3$  progeny means is  $.6441 \left( \frac{.408239}{15} \right) = .017530$  which equals genetic variance of  $F_2$ . Suppose genes are unlinked and have equal effects, and parental values can be used as the genetic range (1.686), Minimum number of genes =  $\frac{(1.686)^2}{8(.017530)} = 20$ , which is an underestimate since parental values need not be the extreme genotypes of the range.

With the variance-ratio method, assuming no epistasy or linkage,

$$\begin{aligned} \text{"Effective" number of genes} &= \frac{\frac{1}{2}(.024531)}{.0003669} = 33, \text{ using } F_3 \text{ data} \\ &= \frac{\frac{1}{4}(.024531) \times 3}{.0002559} = 72, \text{ using } F_4 \text{ data} \end{aligned}$$

As stated before, the variance of variance (.0003669 and .0002559) may not be accurate since the design of experiment has been such that genotypes within progeny are confounded with block variation and the interaction between genotypes and blocks is considerable. Also, the estimation ignores linkage and epistasy, which seem to be operative in this character. If the assumption is not true, an error would be larger with the use of  $F_4$  than  $F_3$  values. Because the average genotypic variance within  $F_4$  progenies is expected to be about one-half of that within  $F_3$  progenies, any deviation from the correct values would be magnified in the  $F_4$  calculation. Therefore, thirty-three is preferred to seventy-two as an estimate of number of genes. However, as the attempt is to calculate the minimum, the actual number may be much larger than thirty-three.

#### D. Dry Weight of Fruit

The five fruits per plant, which had been measured as green weight, were dried and weighed. As a fruit is primarily

composed of varying proportions of seed and fleshy pericarp, which contain different percentages of water and dry matter, the two measurements of weight may be unparallel. But since both green and dry weights are frequently used to express size, it is of interest to determine if the genetical analysis will produce identical results.

High positive correlations between green and dry weights within generations are observed in Table 8. The relationship approximates a linear association. A heavy fruit is expected to contain large amount of dry matter, and vice versa. The small differences between correlation coefficients with arithmetic and logarithmic analyses may be due to the rounding of figures. However, correlation of the mean weights between generations is an exponential one, which is shown in the list below. It is interpreted to mean that a large fruit has a higher percentage of water than a small fruit, which may be one of the features of multiplicative effects of genes.

|                        | <u>P</u> <sub>1</sub> | <u>F</u> <sub>4</sub> | <u>F</u> <sub>3</sub> | <u>F</u> <sub>2</sub> | <u>F</u> <sub>1</sub> | <u>P</u> <sub>2</sub> |
|------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| Green weight mean, gm. | 9.0                   | 53.7                  | 56.3                  | 61.7                  | 59.6                  | 434.9                 |
| Dry weight mean, gm.   | 2.5                   | 8.8                   | 9.1                   | 9.6                   | 10.1                  | 32.2                  |
| Mean log green weight  | 0.948                 | 1.673                 | 1.694                 | 1.758                 | 1.771                 | 2.634                 |
| Mean log dry weight    | 0.388                 | 0.915                 | 0.930                 | 0.963                 | 1.000                 | 1.496                 |

Table 8. Correlation between dry and green weights  
within generations

| Scale of Measurement | P <sub>1</sub> | P <sub>2</sub> | F <sub>1</sub> | F <sub>2</sub> | F <sub>3</sub> | F <sub>4</sub> |
|----------------------|----------------|----------------|----------------|----------------|----------------|----------------|
| Arithmetic           | .84            | .48            | .68            | .89            | .90            | .90            |
| Logarithmic          | .81            | .44            | .68            | .90            | .92            | .91            |

All highly significant at 1% level

The frequencies of dry weight are combined in each generation, and graphical representations drawn in Figures 9 and 10. The data are subjected to a statistical analysis and the estimates given in Tables 9 and 10. Block variance in the logarithmic analysis is at the significant level of 5% probability; so the combining of results of all plots without adjustment introduces some error, which is ignored here.

Dry-weight distributions of segregating generations are positively skewed, and, with logarithmic transformation, they become somewhat negatively skewed. Variances of  $P_1$ ,  $P_2$ , and  $F_1$  fit the logarithmic analysis better than the arithmetic.

The arithmetic  $F_1$  mean (10.1) is lower than the arithmetic average of parents (17.4), but its logarithmic mean (1.000) is higher than the average of parental logarithmic means (0.942). The latter difference implies dominance or epistasy of large, which is confirmed by the decreasing trend of means of  $F_1$ ,  $F_2$ ,  $F_3$ , and  $F_4$  (1.000 to 0.915). Thus, multiplicative gene effects are still exhibited in dry weight. As the means of segregating generations tend to decrease, the positive skewness of the arithmetic distributions cannot be ascribed to dominance of small but to multiplicative gene action; and the negative skewness of logarithmic distributions may be attributed to dominance of large. The trend is not toward the parental average (0.942), so the substrate

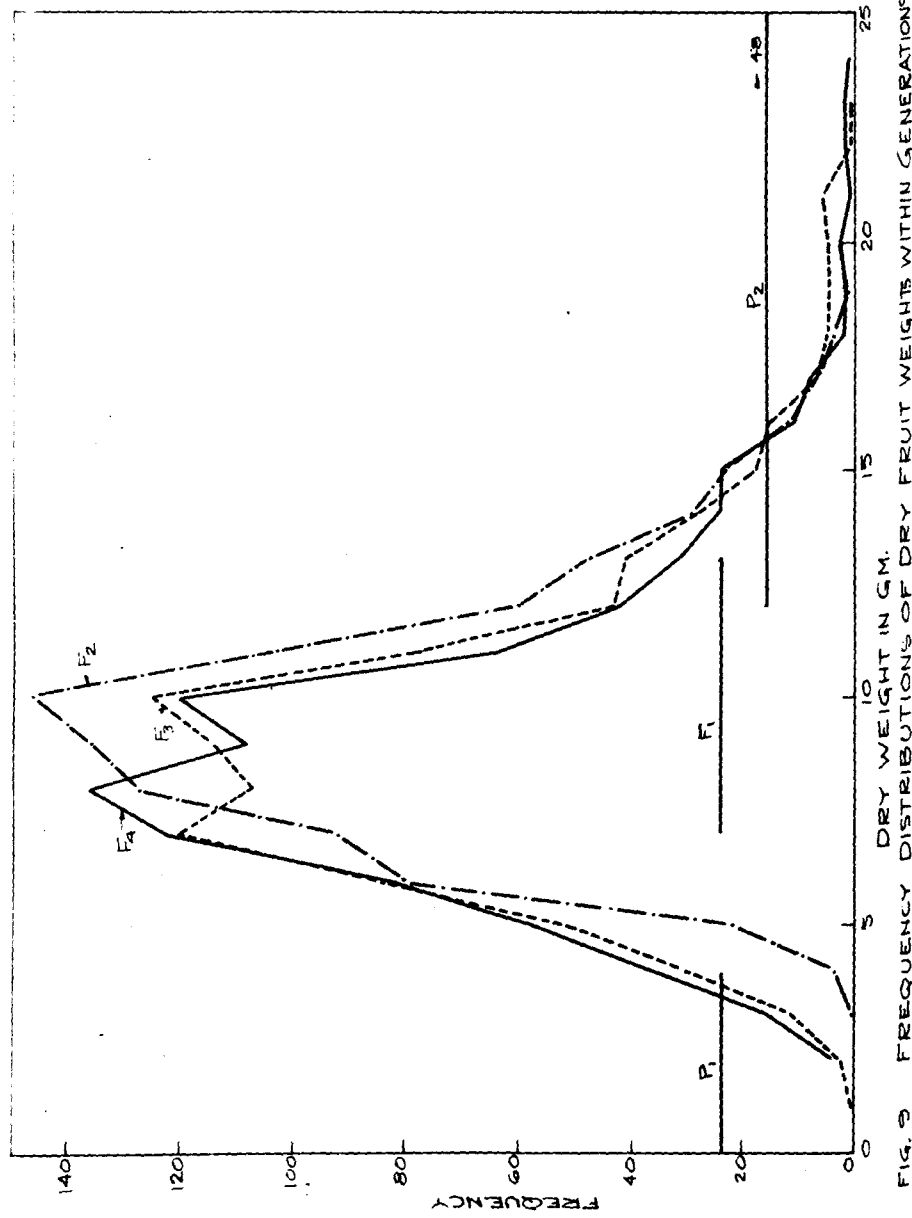


FIG. 9 FREQUENCY DISTRIBUTIONS OF DRY FRUIT WEIGHTS WITHIN GENERATIONS.



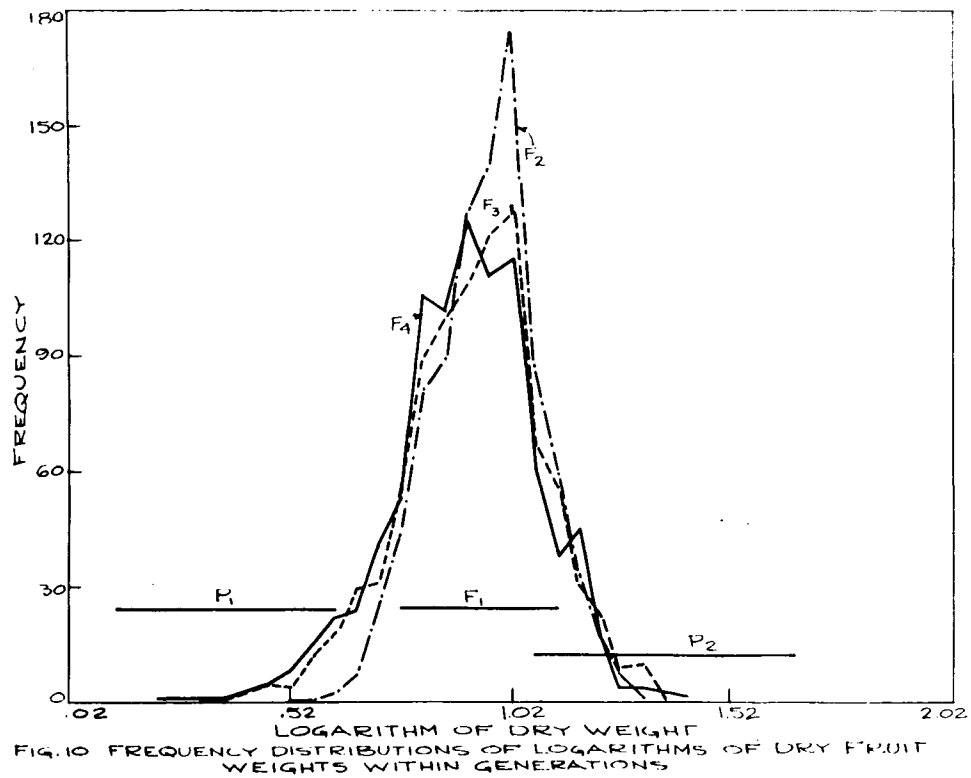


Table 9. Statistics of dry fruit weight,  
arithmetic analysis

|                             | df. | Mean | Skewness <sup>1</sup> | Kurtosis <sup>2</sup> | Mean Square |
|-----------------------------|-----|------|-----------------------|-----------------------|-------------|
|                             |     | gm.  |                       |                       |             |
| Blocks                      | 14  |      |                       |                       | 45.7221     |
| Error                       | 42  |      |                       |                       | 32.0348     |
| P <sub>1</sub> (Red Chili)  | 285 | 2.5  | 0.75**                |                       | 0.2057      |
| P <sub>2</sub> (Sunnybrook) | 283 | 32.2 | -0.11                 |                       | 33.8026     |
| F <sub>1</sub>              | 285 | 10.1 | -0.27                 |                       | 1.1558      |
| F <sub>2</sub>              | 885 | 9.6  | 0.71**                | 0.72**                | 7.2514      |
| F <sub>3</sub> total        | 883 | 9.1  | 0.89**                | 1.44**                | 11.6928     |
| between progenies           | 59  |      |                       |                       | 92.8292     |
| within progenies            | 824 |      |                       |                       | 5.8832      |
| F <sub>4</sub> total        | 882 | 8.8  | 0.98**                | 2.07**                | 10.7539     |
| between progenies           | 59  |      |                       |                       | 77.5175     |
| within progenies            | 823 |      |                       |                       | 5.9677      |

<sup>1</sup>sg<sub>1</sub> =  $\sqrt{\frac{1}{n}}$  .141 for P<sub>1</sub>, P<sub>2</sub> and F<sub>1</sub>; sg<sub>1</sub> =  $\sqrt{\frac{1}{n}}$  .082 for F<sub>2</sub>, F<sub>3</sub>, and F<sub>4</sub>  
<sup>2</sup>sg<sub>2</sub> =  $\sqrt{\frac{1}{n}}$  .163

Table 10. Statistics of dry fruit weight, logarithmic analysis

|                                | df. | Mean  | Skewness<br>$S_1$ | Kurtosis<br>$S_2$ | Mean Square | Genotypic Variance <sup>2</sup> |
|--------------------------------|-----|-------|-------------------|-------------------|-------------|---------------------------------|
| Blocks                         | 14  | 102   |                   |                   | .042471*    |                                 |
| Error                          | 43  |       |                   |                   | .018252     |                                 |
| P <sub>1</sub> (Red Chili)     | 285 | 0.388 | -0.23             |                   | .006077     |                                 |
| P <sub>2</sub> (Sunnybrook)    | 283 | 1.496 | -1.14**           |                   | .006867     |                                 |
| P <sub>1</sub>                 | 285 | 1.000 | -0.84**           |                   | .002523     |                                 |
| P <sub>2</sub>                 | 285 | 0.963 | -0.14             | 0.05              | .014865     | .009708                         |
| P <sub>3</sub> total           | 823 | 0.930 | -0.40**           | 0.61**            | .028057     | .022898                         |
| between progenies <sup>1</sup> | 59  |       |                   |                   | .199332     |                                 |
| within progenies               | 824 |       |                   |                   | .015793     |                                 |
| P <sub>4</sub> total           | 822 | 0.915 | -0.42**           | 0.64**            | .027004     | .021845                         |
| between progenies <sup>1</sup> | 59  |       |                   |                   | .190636     |                                 |
| within progenies               | 823 |       |                   |                   | .015273     |                                 |

<sup>1</sup>Variance between progenies = Variance within progenies + 15(Genotypic variance between progeny means)

<sup>2</sup>Genotypic variance = Total variance - Environmental variance of .005159

Table 11. Estimated variances of dry fruit weight, logarithmic analysis

Average environmental variance = .005159

Regression of  $F_4$  progeny means on  $F_3$  progeny means = .5820

| Generations           | Genotypic Variance     | Genetic Variance       |
|-----------------------|------------------------|------------------------|
| $F_2$                 | .009706                | .007734                |
| $F_3$ total           | .022898                | 1.50(.007734)          |
| between progeny means | .012236                | .007734                |
| within progenies      | $\frac{1}{2}(.009706)$ | $\frac{1}{2}(.007734)$ |
| $F_4$ total           | .021845                | 1.75(.007734)          |
| between progeny means | .011691                | .007734                |
| within progenies      | $\frac{1}{4}(.009706)$ | $\frac{1}{4}(.007734)$ |

Genotypic variance between  $F_3$  progeny means =  $\frac{1}{15}(.199332 - .015793) = .012236$

Genotypic variance between  $F_4$  progeny means =  $\frac{1}{15}(.190636 - .015273) = .011691$

Genetic variance between  $F_3$  progeny means =  $.5820(\frac{.199332}{15}) = .007734$

factors must be modifying the expression of genes in parental populations.

Since the effects of genes and environment are multiplicative, the analysis may then proceed on the logarithmic base. Assuming no epistasy, the excess of dominance of large is:

$$\Sigma d_1 = 4(F_2 \text{ mean} - F_3 \text{ mean}) = 4(.963 - .930) = .132$$

$$\Sigma d_1 = 8(F_3 \text{ mean} - F_4 \text{ mean}) = 8(.930 - .915) = .120$$

Average =  $\frac{1}{2}(.132 + .120) = .126$ , which may be used for prediction of mean dry weight of the succeeding unselected selfed generations.

Variance of  $F_1$  (.002533) is smaller than that of  $P_1$  (.006077) and  $P_2$  (.006867). The average of the three (.005159) is taken as environmental variance for estimating genotypic variances in Table 11. The ratios of genotypic variances are:

$$\frac{H(F_3) - H(F_2)}{H(F_2) - H(F_1)} = \frac{.022898 - .009706}{.009706} = 1.36$$

$$\frac{H(F_4) - H(F_3)}{H(F_3) - H(F_2)} = \frac{.021845 - .022898}{.022898 - .009706} = -.08$$

If there is no linkage or epistatic interaction, and if  $d_1$  is a dominant deviation,

$$\begin{aligned} \text{Variance of } F_2 - \text{Variance of } F_3 \text{ progeny means} &= 3/16 \Sigma d_1^2 \\ &= .009706 - .012236 \end{aligned}$$

$$\Sigma d_1^2 = -.013493$$

$$\begin{aligned} \text{Variance of } F_3 \text{ progeny means} - \text{Variance of } F_4 \text{ progeny} \\ \text{means} &= 3/64 \Sigma d_1^2 \\ \Sigma d_1^2 &= .011627 \end{aligned}$$

The results are identical to those of green weight. It may then be stated that there are indications of linkage in repulsion phase or genetic interaction yet to be identified.

The regression of  $F_4$  on  $F_3$  is .5820. The genetic variance of  $F_2$  is  $.5820(\frac{.199332}{15})$  or .007734. If genes are equal in effects with no linkage, and if genetic range (1.108) is the difference between parental means,

$$\text{Minimum number of genes} = \frac{(1.108)^2}{8(.007734)} = 20$$

Since the individuals within progenies were planted in different plots and block variation is significant, the observed values cannot be used in the calculation. The average variance within progenies is to be estimated from  $F_2$  genotypic variance. If there is no linkage or epistasy,

$$\begin{aligned} \text{Effective number of genes} &= \frac{\frac{1}{2}(.009706)}{.0001488} = 33, \text{ using } F_3 \text{ data} \\ &= \frac{\frac{1}{4}(.009706)}{.00009573} \times 3 = 76, \text{ using} \end{aligned}$$

$F_4$  data.

These numbers are identical to those of green weight.

## V. CONCLUSIONS

### Fruit length and width

Results of statistical analysis favor the hypothesis that mature fruit length and width in the pepper are primarily determined by shape and weight genes. As a consequence, from the breeder's aspect, it will then be impossible to find a plant with fruit length or width beyond the limits of shape manifestation if weight is held constant. Length and width factors, per se, even if they are present, will provide only minor variations in the hybrids after shape effects have been established. The extremely transgressive segregation of length alone can hardly be genetically explained. Correlation studies also lead to the same conclusion that fruit length and width factors are transmitted together as shape. This deduction is in agreement with other investigators on similar problems.

### Fruit shape

The ratio of length and width was used as a shape index. Since environmental variations are multiplicative in  $P_1$ ,  $P_2$ , and  $F_1$ , it is believed, by parallelism, that shape genes also act multiplicatively; thus, the values of shape indices were transformed to logarithms for analysis.

Trimodal distributions and apportionments of frequencies

of  $F_2$ ,  $F_3$ , and  $F_4$  indicate a single gene pair for shape with incomplete dominance of oblate type. Assigning modal values to genotypes, oblate (OO) = .06 (index 1.15), heterozygote (Oo) = .18 (index 1.51), and elongated (oo) = .54 (index 3.43), it is discovered that the total observed genotypic variances are all accounted for by the segregation of this gene pair. The discrepancy between parental types and homozygous genotypes may be ascribed to expressivity of the gene in dissimilar genetic substrates and environmental conditions. With the regression method, heritability of  $F_3$  progeny means was found to be 68 percent for fruit shape.

The one-gene theory confirms the findings by Kaiser (1935) in his ontogenetic studies of the cross IV (oblate) x IX (elongated), in which, on a logarithmic scale, he obtains developmental curves of  $F_2$  fruits: three straight-lined (resulting in oblate shape) to one sickle-shaped (elongated). Inasmuch as Kaiser also gets a reverse ratio in another cross, multiple allelism may be the answer. If the theory is true, then breeding for shape genotypes will be a simple matter. On the other hand, shape manifestation is presumably complicated by genetic substrate factors and environment, which make selection of individual segregants for parental types difficult.

#### Green weight

Using variances and trends of means as criteria, the



environmental variation and genic manifestation of fruit weight were found to be multiplicatively cumulative. The inference is in accord with several investigators who, relying on parental averages, have observed the  $F_1$ ,  $F_2$ , and back-cross values to be closer to geometric than arithmetic means. But a contradiction between the two methods would arise regards dominance relations as is revealed in the results here. The definite decreasing trend of means of segregating generations will lead to a conclusion that genes are dominant for large or epistatic for large even though the  $F_1$  is lower than the average of parental values, from which dominance of small is usually inferred. An argument has been presented that parental and  $F_1$  values are not always dependable for making judgment on properties of genes because genetic substrates in those populations differ from each other.

Weight is the final fruit size attained by growth and development. Weight genes then must be those physiological factors which govern cell division and expansion, of which any variation will be multiplicative. So it is natural to find weight genes manifesting multiplicative action, which, according to recent reports, seem to prevail in many polygenic characters which have been studied. In those characters, a proportionate progress in breeding would be a consequence. Selection for high values would be enhanced by the multiplicative actions of genes.

Results show that 64 percent of variance of  $F_3$  progeny means for fruit weight is genetic. The minimum number of genes for weight differences is of the order of 20 to 33. The prevalent dominance is that of large. With the observed variances, there are indications that some genes are linked in repulsion phase or are epistatic with each other, but no conclusive proof can be given in the present study.

#### Dry weight

Dry weight of fruits is highly correlated with green weight. Genetic analyses of the two give identical results and conclusions with reference to the number and properties of genes. Dry and green weights must be expressions of the same genetic factors. For genetical investigation, it makes no difference whether one chooses to study dry weight or green weight.

It is inferred that multiplicative action of genes is manifested not only in the weight of dry matter of the fruits but also in the water content which increases in percentage as fruits become larger.

## VI. SUMMARY

For the determination of number and nature of polygenes, fruit-size characters, as expressed by length, width, shape index (ratio of length to width), green weight, and dry weight, of the red pepper were chosen. Crosses were made between two varieties, the Red Chili ( $P_1$ ) with small elongated fruit and the Sunnybrook ( $P_2$ ) with large oblate fruit; and, without selection, the hybrids selfed to the fourth filial generation. The parents and all the hybrids were tested in the same year. The data were subjected to statistical analysis and the estimates then compared with genetic expectations based on various types of gene action.

1. Results confirm the supposition that length and width of fruit are largely expressions of shape and weight factors. Evidences show the existence of genes for shape. The presence of genes for length and width per se is unlikely; and even if they are present, they would cause only small deviations after shape manifestations have been accounted for.

2. Environmental variations of shape indices in  $P_1$ ,  $P_2$ , and  $F_1$  populations are multiplicative in nature; probably so are genotypic variations in  $F_2$ ,  $F_3$  and  $F_4$  generations.

Trimodal logarithmic distributions of segregating generations for fruit shape fit a one-gene hypothesis. When modal values are used to evaluate genotypes, the oblate (OO) shape index is 1.15 (log .08), the heterozygote (Oo) 1.51 (log .18), and the elongated (oo) 3.43 (log .54), oblate being partially dominant to elongated. The discrepancy of the homozygous genotypes and the corresponding parental types is attributed to genetic substrate dissimilarities and environmental modifications.

The one-gene theory closely agrees with Kaiser's conclusion (1935) drawn from his studies of pepper fruit development. The numbers of genes, five and fifty-five, estimated from Castle and Wright's formula and Panse's variance ratio respectively, seem improbably high.

3. For green weight, the effects of genes and environment are found to be multiplicative. Heritability of  $F_3$  progeny means is 64 percent.

The minimum number of genes in this pepper cross has been estimated to be between twenty and thirty-three. The decreasing trend of means of selfed generations suggests that the genes are preponderantly dominant for large or epistatic for large weight. Substrate factors seem to affect expression of the genes. A study of variance reveals some evidences of linkage and epistasy.

4. Dry weight and green weight are strongly correlated. Both must be expressions of the same genes since identical results are obtained from the genetic analyses. Heritability of  $F_3$  progeny means for dry weight is 58 percent.

## VII. ACKNOWLEDGEMENT

The author gratefully acknowledges the encouragement, advice, and help given by Dr. E. W. Lindstrom, under whom this investigation was conducted.

VIII. LITERATURE CITED

- Castle, W. E. 1921. An improved method of estimating the number of genetic factors concerned in cases of blending inheritance. *Sci.* 54: 223.
- Charles, D. R. and H. W. Smith. 1939. Distinguishing between two types of gene action in quantitative inheritance. *Genetics* 24: 34-48.
- Dale, E. E. 1929. Inheritance of fruit length in Capsicum. *Mich. Acad. Sci. Arts and Letters Papers* 9: 89-109.
- De Haan, H. 1931. Contributions to the genetics of Pisum. 'S-Gravenhage, Martinus Nijhoff. p. 321-439.
- Deshpande, R. B. 1933. Studies on Indian Chilis. (3) The inheritance of some characters in Capsicum annum L.. *Indian Jour. Agr. Sci.* 3: 219-300.
- Fisher, R. A., F. R. Immer, and O. Tedin. 1932. The genetical interpretation of statistics of the third degree in the study of quantitative inheritance. *Genetics* 17: 107-124.
- Gustafsson, A. 1946. The effect of heterozygosity on variability and vigor. *Hereditas* 32: 263-286.
- Halsted, B. D. 1915. Inheritance in peppers. *N. J. Agr. Exp. Sta. Ann. Rpt.* 36: 272-283.
- Hutchins, A. E. 1934. Morphological relationships in the ontogeny of the cultivated cucumber, Cucumis sativus L.. *Minn. Agr. Exp. Sta. Tech. Bul.* 96. 35 p.
- Kaiser, S. 1935. The factors governing shape and size in Capsicum fruits; a genetic and developmental analysis. *Torrey Bot. Club Bul.* 62: 433-454.
- Khambanonda, I. 1941. Fruit size inheritance in Capsicum varieties. Unpublished M.S. thesis. Library, Iowa State College, Ames, Iowa.
- Lindstrom, E. W. 1927a. Linkage of size, shape, and color genes in Lycopersicum. *Verh. V. Internat. Kong. Vererb. Berlin.* 2:1031-1057.

- Lindstrom, E. W. 1927b. The inheritance of ovate and related shapes of tomato fruits. Jour. Agr. Res. 34: 961-985.
- Lindstrom, E. W. 1948. The inbreeding 'freeze' of linked genes and its relation to internal chromosome balance of polygenes. Genetics 33: 114. (Abstract).
- MacArthur, J. W. and L. Butler. 1938. Size inheritance and geometric growth processes in the tomato fruit. Genetics 23: 253-268.
- Mather, K. 1943. Polygenic inheritance and natural selection. Biol. Rev. Cambridge Phil. Soc. 18: 32-64.
- Pansey, V. G. 1940. Application of genetics to plant breeding. II The inheritance of quantitative characters and plant breeding. Jour. Genet. 40: 283-302.
- Poole, C. F. and P. C. Grimball. 1945. Interaction of sex, shape, and weight genes in watermelon. Jour. Agr. Res. 71: 533-552.
- Powers, LeRoy. 1934. The nature and interaction of genes differentiating habit of growth in a cross between varieties of Triticum vulgare. Jour. Agr. Res. 49: 573-605.
- Powers, LeRoy. 1941. Inheritance of quantitative characters in crosses involving two species of Lycopersicon. Jour. Agr. Res. 63: 149-174.
- Powers, LeRoy. 1942. The nature of the series of environmental variances and the geometric means in crosses involving species of Lycopersicon. Genetics 27: 561-575.
- Sax, K. 1924. The nature of size inheritance. Nat. Acad. Sci. Proc. 10: 224-227.
- Shull, G. H. 1921. Estimating the number of genetic factors concerned in blending inheritance. Amer. Nat. 55: 556-564.
- Sinnott, E. W. and C. B. Durham. 1929. Developmental history of the fruit in lines of Cucurbita pepo differing in fruit shape. Bot. Gaz. 87: 411-421.
- Sinnott, E. W. 1935. Evidence for the existence of genes controlling shape. Genetics 20: 12-21.



- Straus, F. S. 1942. A genetic mechanism of heterosis as demonstrated by egg production in Drosophila. Unpublished Ph.D. thesis. Library, Iowa State College, Ames, Iowa.
- 'Student'. 1934. A calculation of the minimum number of genes in Winter's selection experiment. Ann. Eugenics 6: 77-82.
- Webber, H. J. 1912. Preliminary notes on pepper hybrids. Amer. Breed. Assn. Ann. Rpt. 7: 188-199.
- Weetman, L. M. 1935. Inheritance and correlation of shape, size, and color in the watermelon. Iowa Agr. Exp. Sta. Res. Bul. 228: 222-256.
- Wright, S. 1922. The effects of inbreeding and crossbreeding on guinea pigs. U. S. Dept. Agr. Bul. 1121. 47 p.
- Wright, S. 1934. The results of crosses between inbred strains of guinea pigs, differing in number of digits. Genetics 19: 537-551.
- Yeager, A. F. 1937. Studies on the inheritance and development of fruit size and shape in the tomato. Jour. Agr. Res. 55: 141-152.